Title
The Structure and Activation of Soluble Guanylate Cyclase

Author
Horst, Benjamin Guthrie

Publication Date
2019

Peer reviewed|Thesis/dissertation

The Structure and Activation of Soluble Guanylate Cyclase

By

Benjamin G. Horst

A dissertation submitted in partial satisfaction of the

requirements for the degree of

Doctor of Philosophy

in

Chemistry

in the

Graduate Division

of the

University of California, Berkeley

Committee in charge:
Professor Michael A. Marletta, Chair
Professor Judith P. Klinman
Professor Michelle C. Y. Chang
Professor Michiko E. Taga

Fall 2019

The Structure and Activation of Soluble Guanylate Cyclase

Abstract

The Structure and Activation of Soluble Guanylate Cyclase

by

Benjamin G. Horst

Doctor of Philosophy in Chemistry

University of California, Berkeley

Professor Michael A. Marletta, Chair

The biochemical signaling pathways that involve gaseous primary signaling molecules have only been characterized in the last three decades, due to the small diversity of diatomic gases and the lack of other known physiological functions of these gases (apart from O_2 as the terminal electron acceptor in oxidative phosphorylation). In general, certain events must take place for any process to be considered biochemical signaling: the primary signaling molecule must be synthesized and excreted from the generator cell or be present or absent in the environment, that signaling molecule must come in contact with its specific receptor molecule within the receiving cell, and the receptor molecule must transduce the primary singling molecule into a secondary signal which activates downstream signaling pathways. Only a pathway that is comprised of signal generation, the primary signal, and signal reception constitutes a complete biochemical signaling pathway. In spite of the relatively recent discovery, gaseous signaling molecules control critical signaling pathways in many aspects of organismal physiology.

In the 1980s, two different lines of research serendipitously came together to describe the first gaseous biochemical signaling pathway. The first was pharmacologists' observation that vasculature that had the endothelium removed would constrict, and thus there must be an Endothelium Derived Relaxation Factor (EDRF) that maintained appropriate vasculature tone. Curiously, EDRF had the exact same chemical properties as nitric oxide (NO), however there was no known biotic source of this toxic gas. The second line of research centered on the observation that excreted nitrates spiked in concentration when sickness was incurred, even when nitrate intake remained the same. This led to the discovery that macrophages synthesize the cytotoxic agent NO to attack invading pathogens, and thus a biotic source of NO had been discovered. The EDRF and NO were shown to be one and the same, and this coupled with the fact that soluble guanylate cyclase had been known to be stimulated by NO since the 1970s provided the receptor of the first gaseous biochemical signaling pathway.

The enzyme responsible for the generation of NO is nitric oxide synthase (NOS) (EC 1.14.13.39). NOS catalyzes the 5-electron oxidation of arginine to citrulline and NO. There are three different isoforms of NOS in Homo sapiens, inducible NOS (iNOS), endothelial NOS

(eNOS), and neuronal (nNOS). Nitric oxide synthases are activated by calcium-bound calmodulin (Ca^{2+}-CaM). iNOS has a high affinity for Ca^{2+}-CaM, and thus is regulated at the transcriptional level in macrophages where it synthesizes NO at sites of infection or inflammation. eNOS and nNOS have lower affinity for Ca^{2+}-CaM and are constitutively expressed, which allows calcium gradients in neurons and endothelial cells to control the activation of these isoforms. For example, when endothelial cells hyperpolarize due to high blood pressure, intracellular concentrations of Ca^{2+} increase up to ten-fold which in turn increases Ca^{2+}-CaM concentration and thus activates eNOS.

Activated eNOS generates NO at nanomolar concentrations, which diffuses to nearby smooth muscle cells that express the mammalian NO receptor, soluble guanylate cyclase (sGC) (EC 4.6.1.2). sGC is a heterodimeric hemeoprotein which can sense NO at nanomolar concentrations. When NO binds to sGC, the enzyme increases its catalytic activity of cyclizing 5′-guanosine triphosphate (GTP) to 3′,5′-cyclic guanosine monophosphate (cGMP). The molecular mechanisms by which NO activates sGC are not completely understood. In particular, it is clear from experiments both *in vitro* and in cells that one equivalent of NO does not activate sGC to its maximal activity but only to a low activity state. Excess NO must be present in order to completely stimulate the enzyme. If that excess NO is removed either with a buffer exchange or a chemical trap, the enzyme returns to a low activity state. The increase in cellular concentration of cGMP activates a variety of downstream signaling proteins, including cGMP-dependent protein kinases, cGMP-regulated ion channels, and phosphodiesterases that mediate the downstream cell signaling cascade. The resultant phenotype of these cellular signaling molecules is the relaxation of the smooth muscle cells, the dilation of the vasculature, and the decrease of overall blood pressure. Thus, any malfunctions of this pathway can result in too high of blood pressure and increased risk of cardiovascular diseases.

The NOS-NO-sGC-cGMP signaling pathway has been the target for stimulating therapeutics to treat forms of high blood pressure. A screen of small molecules that induce platelet aggregation found that the chemically synthesized benzylindazole YC-1 was able to stimulate sGC in a heme-dependent manner. Significant efforts in medicinal chemistry in the following two decades yielded riociguat, the first FDA-approved sGC stimulator that is used to treat certain forms of pulmonary hypertension. However, the specific binding site of YC-1 derived compounds on sGC and the reason for the specificity of these types of molecules against the soluble form of guanylate cyclases has not been elucidated. Additionally, atypical soluble guanylate cyclases have been described that sense O_2 instead of NO. Thus, the possibility exists for gases other than NO and O_2 to function as signaling molecules in eukaryotic organisms. These outstanding questions will be answered in the following thesis entitled: The Structure and Activation of Soluble Guanylate Cyclase.

TABLE OF CONTENTS

LIST OF ABBREVIATIONS

1-NO	one equivalent of NO
5c	five coordinate
6c	six coordinate
AC	adenylate cyclase
ABPP	activity-based protein profiling
AEBSF	4-(2-aminoethyl)benzenesulfonyl fluoride hydrochloride
BCA	bicinchoninic acid assay
BIC	butyl isocyanate
BSA	bovine serum albumin
CAT	catalytic
Cα	alpha carbon
CC	coiled-coil
cDNA	complementary DNA
CO	carbon monoxide
Cr	*Chlamydomonas reinhardtii*
cGMP	3′,5′-cyclic guanosine monophosphate
Cre	*Chlamydomonas reinhardtii*
Cryo-EM	cryoelectron microscopy
CV	column volume
Da	dalton
DEA NONOate	diethylammonium (Z)-1-(N,N-diethylamino)diazen-1-ium-1,2-diolate
DEER	double electron-electron repulsion
DETA NONOate	(Z)-1-[N-(2-aminoethyl)-N-(2-ammonioethyl)amino]diazen-1-ium-1,2-diolate
DMSO	dimethyl sulfoxide
DTT	dithiothreitol
EPR	electron paramagnetic resonance
FPKM	fragments per kilobase of transcript per million mapped reads
FRET	Förster Resonance Energy Transfer
GSNO	*S*-nitroso glutathione
GTP	5′-guanosine triphosphate
H-NOX	heme nitric oxide/oxygen binding
H/D	hydrogen/deuterium
HEPES	4-(2-hydroxyethyl)piperazine-1-ethanesulfonic acid
HS Media	high salt growth media (used for photosynthetic growth of *C. reinhardtii*)
HDX-MS	hydrogen deuterium exchange mass spectrometry
HNOB	heme nitric oxide/oxygen binding
HPLC	high pressure liquid chromatography
HUVEC	human umbilical vein endothelial cells
IAC	iodoacetic acid
IAM	iodoacetamide
IAyne	N-(5-hexyn-1-yl)acetamide
IMAC	immobilized metal affinity chromatography
IsaAnh	isatoic anhydride

IPTG	isopropyl β-D-1-thiogalactopyranoside
JGI Phytozome	The Joint Genome Institute Photozome webpage, a Department of Energy Office of Science User Facility (https://phytozome.jgi.doe.gov/pz/portal.html)
Maleimide-yne	1-(5-Hexyn-1-yl)-1H-pyrrole-2,5-dione
MALS	multi-angle light scattering
MMTS	*S*-methyl methanethiosulfonate
Ms	*Manduca sexta*
MW	molecular weight
MS	mass spectrometer/mass spectrometry
MWCO	molecular weight cut off
NEM	N-ethyl maleimide
NO	nitric oxide
NOS	nitric oxide synthase
O_2	oxygen
P(r)	pair distance function
PAS	Per/Arnt/Sim
PCR	polymerase chain reaction
PDB	Protein Data Bank (http://www.rcsb.org)
PROLI NONOate	1-(hydroxy-NNO-azoxy)-L-proline, disodium salt
RFQ	rapid freeze quench
R_g	radius of gyration
Rn	*Rattus norvegicus*
RT-qPCR	reverse transcriptase mediated quantitative polymerase chain reaction
S-helix	signaling helix
s-NHS-Ac	sulfosuccinimidyl acetate
SAXS	small angle X-ray scattering
SDS-PAGE	sodium dodecyl sulfate polyacrylamide gel electrophoresis
SEC-MALS	size-exclusion chromatography in line with multi-angle light scattering
SEC-SAXS	size-exclusion chromatography in line with small angle X-ray scattering
SF9	*Spodoptera frugiperda* clonal isolate
sGC	soluble guanylate cyclase
So	*Shewanella oneidensis*
TAP	tris, acetate, phosphate growth media (used for mixotrophic and dark growth of *C. reinhardtii*)
TEA	triethanolamine
TOF	time of flight
U	unliganded
UTR	untranslated region
UV-vis	ultra violet visible
xsNO	excess nitric oxide

LIST OF FIGURES

LIST OF TABLES

ACKNOWLEDGEMENTS

There are many people that I need to thank for helping me complete my educational journey over the past five and a half years. I feel extremely confident in saying that I would not be where I am today without each and every one of you.

Firstly, to my principle investigator, scientific inspiration, and colleague Michael A. Marletta. Thank you for sharing your countless years of experience and wisdom with a young, excited chemist who finally settled on the idea of switching to biochemistry after exploring biophysics, inorganic chemistry, bioinorganic chemistry, and synthetic biology. I don't think I've told you this story, but I first came across your lab as a senior undergraduate at UW in the spring of 2013. I was sitting in the Chemistry Study Center there and pulling papers from labs that looked interesting as I was beginning to look at graduate schools. I read Will Beason's *JACS* communication about PMOs and Brian Smith's *Biochemistry* article about S-nitrosation of NOS. I thought that someone who was thinking as chemically about proteins as you were might be someone that I would want to work for. I was quite disappointed to find out in the following months that you had moved to La Jolla just after those papers had come out. You do know how the rest of the story turned out, but I did want to thank you here for answering the email I sent in late April of 2015 asking if you might be willing to take on a fourth rotation student immediately upon your return to Berkeley. We met at Yali's at the beginning of May 2015 and I told you that I thought it was fascinating that NO as a diatomic gas could activate such a large protein as sGC (which is still true to this day). The next few years have absolutely flown by.

I have learned a great deal about science, teaching, and mentorship from you over the past four and a half years. We have discussed at great length the perils of deviating too far from the natural world in our reductionist biochemical approach, whether it is in terms of isolation protein domains or applying non-physiological amounts of molecules to a system. I enjoyed teaching with you in Chemical Biology and learned the importance of communicating the history of the science as well as the power of a well-timed joke every now and then. Finally, I have very much enjoyed observing your mentorship style both individually and to the lab as a whole. Your teachings will continue through me, as they do through many of your former mentees.

To my lab mates, thank you for sharing your time, your past experiences, and your outlook on life with me over the years. I have enjoyed working with all of you and I appreciate your willingness to both teach and learn from me as I've grown over the years. To Charles Hespen PhD, Minxi Rao PhD, Elise Span PhD, Elizabeth Ndontsa PhD, Joel Bruegger PhD, Yirui Guo PhD, John Hangasky III PhD, Marco Agostoni PhD, Tracy Hinder, Megan Brophy PhD, Tyler Detomasi, Aren Nazarian, Angela Yang, Christopher Lemon PhD, Stefan Kapczynski, Matthew Cooper, Edna Stewart, Elizabeth Wittenborn PhD, Silver Alkhafaji, Vanha Pham, Skylar Henry, Sarah Chen, Erika Wirachman, Miriam Hood, Molly Hydron, Kimberly Houghton, Allison Batka, Alejandra Martinez PhD, and Kristian Bolanos Ibarra, thank you for being excellent lab mates and the wonderful discussions about our science over the years. I want to briefly highlight the people who were instrumental in helping me start and finish my first project, Charles Hespen PhD and John Hangasky III PhD. I also want to thank the postdocs who helped edit countless presentations and publications, Megan Brophy PhD, Christopher Lemon PhD, and Elizabeth Wittenborn PhD. I

feel that even though I will receive my blue and gold PhD from Berkeley, there will be a tinge of MIT cardinal red thrown in there as well.

I want to especially thank my four direct mentees Aren Nazarian, Edna Stewart, Miriam Hood, and Kimberly Houghton for being patient with me as I learned how to mentor each of you. These four relationships have been one of the most treasured parts of my entire graduate school experience and I hope to continue in a mentorship role for many years to come. Thank you for bearing with my strict documentation schemes and my ever-present ways as we progressed along different project related to sGC. It was a pleasure to watch you present the work we have done together. I hope I was able to guide you in a positive direction throughout your time in the Marletta lab. Keep in touch, I would love to hear how everything is progressing in 6 months, a year, five years, and beyond.

If there is one lesson that I have learned during my PhD, it is that science cannot be done in isolation. We are experts in such a small slice of the scientific endeavor that to not ask for help would be foolhardy. Therefore, I would like to extend my gratitude to all of the scientific collaborators that I have worked with over the years. To Adam Yokom PhD, Kyle Morris PhD, and James Hurley PhD, thank you for giving *Ms* sGC your time and effort over the course of two years, I am proud that we were able to propose a binding pocket for the precursor to an FDA-approved drug that had not been discovered before! To Daniel Rosenberg and Michal Hammel PhD, thank you for allowing me to commandeer a significant portion of the SEC-SAXS time during the last year. I hope there will be more SAXS performed on sGC at the ALS in the near future. There have also been collaborations in which the science did not work out how we expected. Thus, I would like to thank Carl Ward, Daniel Nomura PhD, Brian Sanders PhD, Harry Gray PhD, Georgios Katsoukis PhD, and Heinz Frei PhD for their willingness to spend some of your time in collaborations that proved to be more challenging than we imagined at the start.

Next, I would like to thank the members of both my qualification committee and my thesis committee, Judith Klinman PhD, Michelle Chang PhD, Anne Baranger PhD, and Michiko Taga PhD. I am proud that (apart from Michael), my entire committee is made up of female professors that have either created or are leading their respective fields. Although we have a long way to go towards gender (and other) equality in academia, I was thrilled to have had your input into my scientific path. Your feedback during my qualifying exam was something that I took to heart as I continued my scientific journey. And I especially want to thank you for your comments during our committee meeting, which were instrumental in building my confidence as I neared the end of this educational journey.

I have also had many scientific mentors and colleagues who have aided me as I transitioned into the field of biochemistry and attempted to learn new techniques that I did not, or my lab mates did not, know. My three rotation mentors Joseph Cotruvo PhD, Kevin Metcalf PhD, and Ningkun Wang PhD brought me up to speed on basic molecular biology and biochemistry that I needed to learn. Apart from Cyg11, all of the other sGCs were expressed in insect cells which I learned to culture from Xiaoxian Cao from John Kuriyan's lab and Ursula Schulze-Gahmen PhD from Jim Hurley's lab. Finally, the Niyogi lab members Erika Erikson PhD, Haruhiko Jimbo, and Setsuko Wakao PhD were always willing to speak with me about *Chlamydomonas reinhardtii* growth, culture, and transcriptomics and I was very appreciative of their time.

Coming to UC Berkeley affords one with approximately seventy colleagues who are all at the same stage of their careers and more than willing to lend support to one another, especially during the early years of the PhD. From both the first-year chemists and the MCB members of the Chemical Biology Graduate program, I have made lifelong friends out of a few of my cohort mates, and many others will continue to serve as colleagues for years to come. I am also very lucky to be surrounded by many passionate teachers. To all of the GSIs that I have taught with, instructors of the QB3 Bootcamp, and members of the GSI Teaching and Learning center, thank you for your hard work and support over the years.

I am fortunate to have worked in Stanley Hall over the past four and a half years, and have sincerely enjoyed interacting with Harry Stark, Thom Opal, Dave Rodgers, Donna Hendricks, and Kris Thompson. They always did their best to make sure the behind the scenes tasks were taken care of so that we could focus on the science. Also, the receiving team Mike, Chris, Martin, and the VWR Stockroom lead Kevin have been so fun to get to know over the years and are always ready to talk sports when a mental break from science is required. I also would be amiss without thanking our amazing lab custodian Dawn. Thank you for all that you do, and I will miss seeing your always-smiling face after a long evening of work.

To Berkeley's Counseling and Psychological Services, thank you for providing a consistent safety net of mental support that was required to successfully navigate through graduate school and life in general. If you are at all struggling mentally with issues big or small, please seek out CPS or a mental health care professional. You do not have to go this alone.

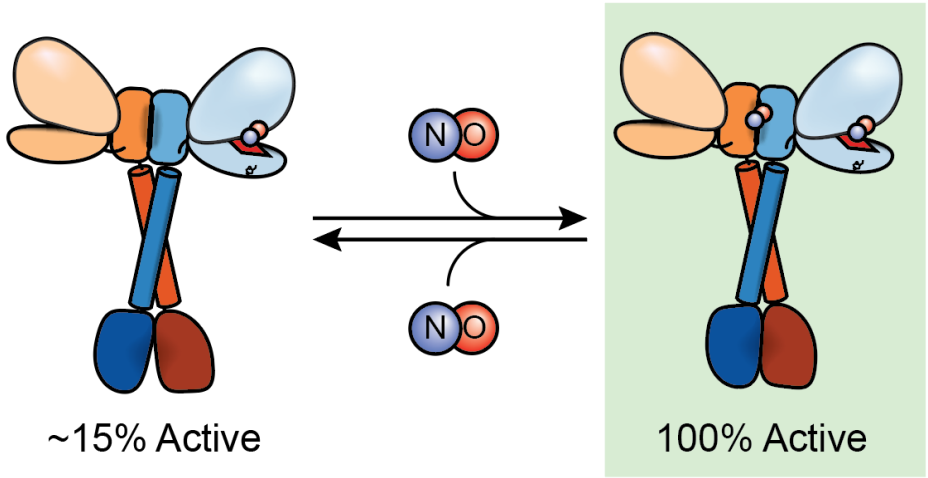
And finally, I want to thank my family. To my parents Arlen Horst and Kathy Macdonald, thank you for always allowing me the space and resources to pursue whatever I seem to be most interested in at a given moment in time. To my siblings Martin and Shannon Horst, thank you for being constant supporters and continual inspirations during my time at Berkeley. And finally, I would like to thank my wife, Emily Guthrie. Your unwavering support of me over the past five years has meant the world to me. From helping me to build confidence in my writing to our whiteboard discussions about our projects; from working through how to teach together when we were the only two GSIs in Chemical Biology to helping me troubleshoot my science problems by asking me how I would advice a younger student who was had a similar issue. This PhD would have been completely different without you, and I'm so happy that you were by my side to shape my experiences for the better.

CHAPTER 1

PHYSIOLOGICAL ACTIVATION AND DEACTIVATION OF SOLUBLE GUANYLATE CYCLASE

Summary

Soluble guanylate cyclase (sGC) is responsible for transducing gaseous signaling molecules into the ubiquitous secondary signaling messenger cyclic guanosine monophosphate in eukaryotic organisms. Mammalian sGC is exquisitely tuned to sense low levels of nitric oxide (NO), allowing cells to respond to non-toxic levels of NO. Other eukaryotic sGCs respond to oxygen (O_2) to regulate motility and behavior, while other gases await the discovery of protein-specific receptors. In this chapter, sGC is discussed in the context of its activation and deactivation in response to gas molecules. The sequence of events in the activation pathway are developed into a comprehensive model of *in vivo* sGC activation as elucidated both from studies with purified enzyme and those done in cells. This model is then used to discuss the deactivation of sGC, as well as the molecular mechanisms of pathophysiological deactivation.



Portions of this chapter have been reported in a previous publication.[1]

Introduction

In the 1980s, nitric oxide (NO) was first characterized as critical to both innate immunity and endogenous signaling in animals.[2–6] NO was the first gaseous signaling molecule synthesized by animals to have its biochemical signaling pathway fully described.[3] Physiologically, NO signaling causes relaxation of vascular smooth muscle, inhibition of platelet aggregation in the vasculature, and modulation of various forms of neurotransmission.[7,8] Beyond these well-established functions, additional aspects of NO signaling continue to emerge, including in insect development and sensory systems,[9] as well as in human pathologies, such as early-onset achalasia[10] and cancer proliferation.[11]

Soluble guanylate cyclase (sGC), a eukaryotic nitric oxide receptor, is a central component in NO-dependent signaling.[4,12] sGC converts 5′-guanosine triphosphate (GTP) to 3′,5′-cyclic guanosine monophosphate (cGMP). When NO binds to sGC, enzyme activity increases several hundredfold, transducing the gaseous paracrine-signaling molecule to a ubiquitous secondary signaling molecule. The molecular details of how NO activates sGC to maximal physiological activity are not fully understood, though much progress has been made.

Aberrations resulting in decreased function of this sGC-dependent signaling pathway have been linked to multiple pathologies, including cardiovascular disease, hypertension, asthma, and neurodegeneration.[13,14] Many emergent pharmaceuticals seek to increase sGC activity in these diseased states. One set of examples are the sGC stimulators, initially exemplified by the benzylindazol compound YC-1[15] and culminating in Bayer's riociguat (Adempas®), the latter was approved by the FDA in 2013 to treat pulmonary arterial hypertension and chronic thromboembolic pulmonary hypertension.[16] This class of molecules has also added to our understanding of the NO-dependent activation mechanism of sGC, however it is still unclear how these molecules by themselves promote the active conformation of sGC[17,18].

This chapter combines structural and functional aspects of NO-dependent activation and deactivation of sGC to build a model of sGC activation *in vivo*. Stimulators and activators are discussed in the context of physiological activation, and interested readers are directed to other reviews[18,19] that are dedicated to these small molecules.

Structure of Soluble Guanylate Cyclase

Soluble guanylate cyclase is a heterodimer comprised of an α and a β subunit (Figure 1.1A). Multiple isoforms of these two subunits have been identified, yet sGC is most commonly expressed as a heterodimer composed of α_1 and β_1 subunits (Uniprot IDs of *Homo sapiens* proteins: Q02108 and Q02153).[20] Each subunit consists of four domains: a heme nitric oxide and oxygen binding (H-NOX) domain,[21] a Per/Arnt/Sim (PAS)-like domain, a coiled-coil (CC) domain, and one subunit of the catalytic heterodimer (CAT). Although the H-NOX subunit structures are similar, only the β subunit has the capacity to bind heme. Since a high-resolution structure of full-length sGC has only been obtained in the last few months, in the past five years the vast majority of structural information has been derived from isolated domains of sGCs. Others have recently

reviewed the structure of sGC,[22,23] and thus the discussion here will be limited to pertinent details and recent developments.

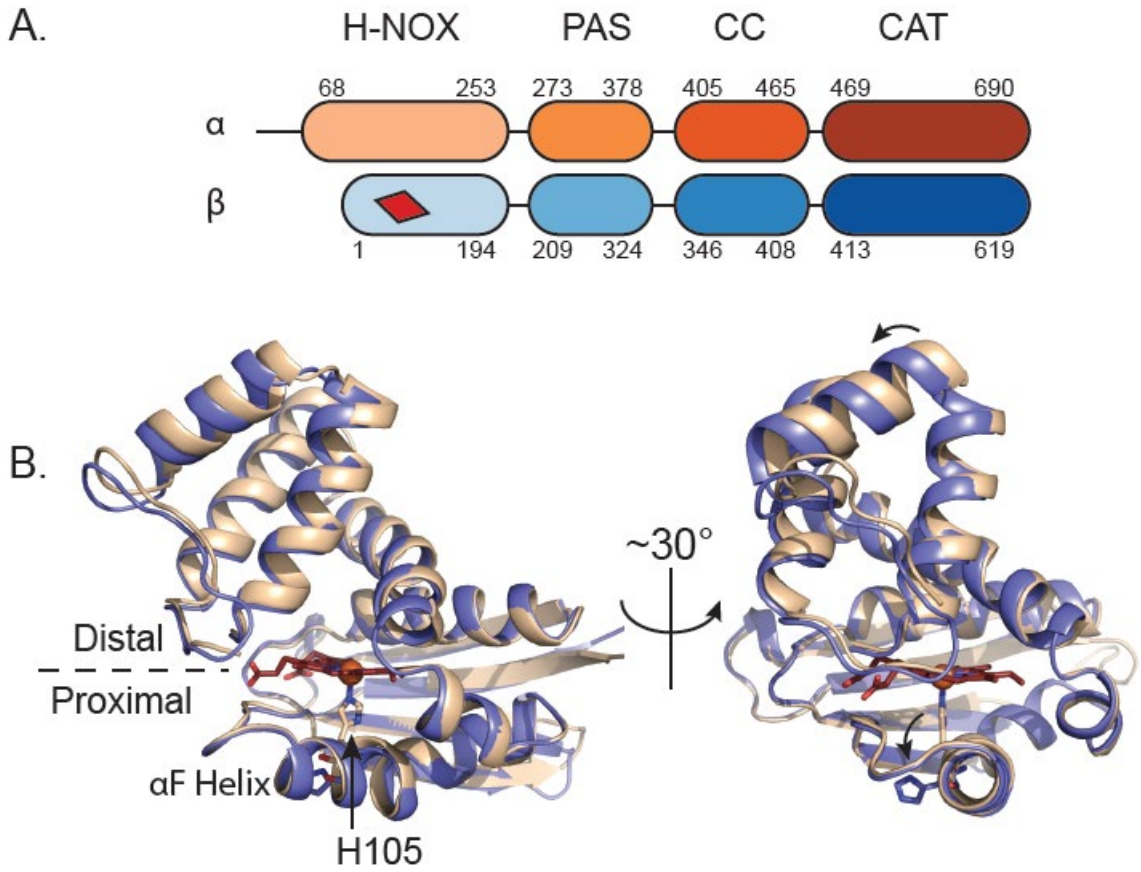


Figure 1.1 – Domain and subdomain organization of sGC A) Schematic of soluble guanylate cyclase domains. The four domains of the α (in orange) and β (in blue) subunits are listed above: heme nitric oxide and oxygen binding domain (H-NOX), Per-Arnt-Sim domain (PAS), coiled-coil domain (CC), and the catalytic domain (CAT). The approximate numbering of the Human α1 and β1 domains is shown. The heme cofactor, shown as a red diamond, binds to β1 H105.
B) Motions of heme nitric oxide and oxygen binding domains (H-NOX) upon nitric oxide binding. Overlay of the *Shewanella oneidensis* H-NOX domain in the Fe^{2+} (wheat, PDB ID: 4U99) and Fe^{2+}–NO (blue, PDB ID: 4U9B) states. Structures have been aligned by the proximal domains (residues 95–180). Displacement of the distal domain and rotation of the proximal histidine are noted with arrows.

H-NOX domains, also known as heme nitric oxide binding (HNOB) or sensor of nitric oxide (SONO) domains,[24] consist of an N-terminal alpha helical subdomain and a C-terminal subdomain composed of alpha helices and four antiparallel beta sheets. A single heme cofactor binds in a central cavity between the two subdomains (Figure 1.1B). The N- and C- terminal subdomains are termed distal and proximal, respectively, with the proximal subdomain defined by the presence of a heme-ligating histidine residue (termed the proximal histidine). H-NOX domains are also found as stand-alone proteins in prokaryotes,[25] and the study of these proteins has greatly improved our understanding of how these domains bind gaseous ligands and then transduce chemical signals.[26–29] These studies, as well as spectroscopic studies of sGC, showed that NO

initially binds to the 5-coordinate (5c) high-spin ferrous heme cofactor (5c His–Fe^{2+}, see Figure 1.2), which leads to a transient 6-coordinate (6c) complex (6c His–Fe^{2+}–NO).[30–32] This complex, driven by the strong σ trans effect NO exerts on the proximal histidine, spontaneously dissociates to the 5c low-spin ferrous nitrosyl complex (5c Fe^{2+}–NO).[33] In turn, cleavage of the iron–histidine bond results in a 90° rotation of the α-helix (αF) that contains the histidine; this conformational change is thought to be the initial step in propagating the NO signal.[34] Some studies have examined the next steps in signal transduction pathway by examining other residues on the αF helix, however, it remains unclear exactly how these events confer full activity to the enzyme.[35,36]

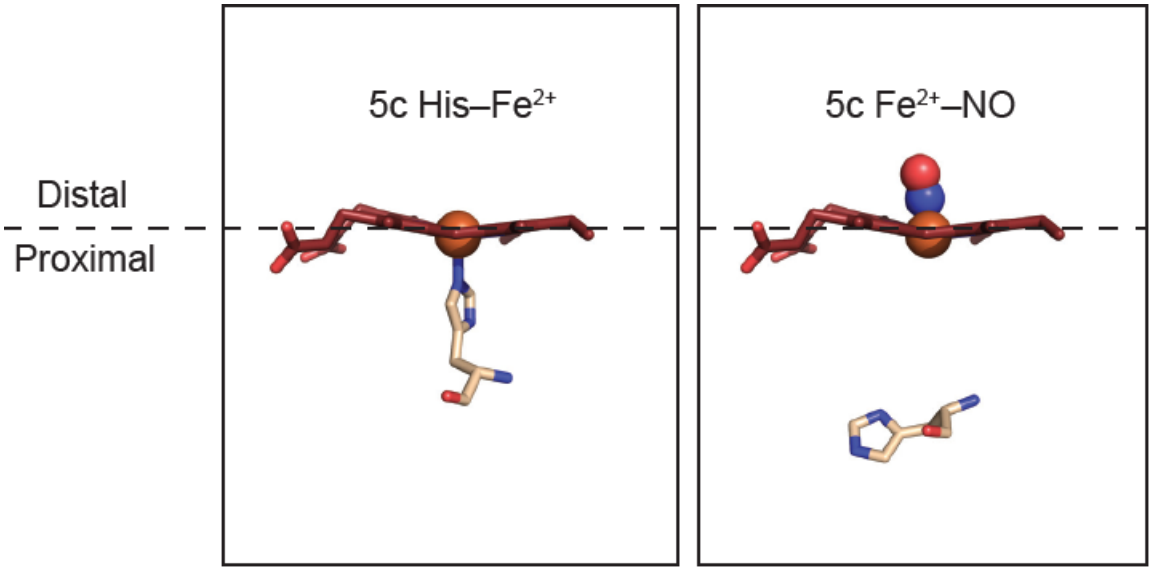


Figure 1.2 – The heme cofactor in sGC has an inherent sidedness as defined by the presence of the coordinating histidine residue (H105 in humans). The histidine binds to the ferrous iron of the heme cofactor from one side, termed the proximal side, as shown. The opposing side of the heme is referred to as the distal side. Upon binding of NO to the distal side of the heme, the strong σ trans effect induces rupture of the iron–histidyl bond, which is thought to be the first step in signal transduction. However, this cleavage event results in the opening of the proximal coordination site for an additional NO molecule to bind. To clarify the potential heme ligation states in sGC, the following naming convention will be used in this review: 4/5/6c (proximal ligand)–Fe^{2+}–(distal ligand), where 4/5/6c refers to the coordination number of the Fe^{2+} center. For example, unliganded ferrous sGC with the proximal histidine bound would be described as 5c His–Fe^{2+}, whereas a distal ferrous nitrosyl complex would be termed 5c Fe^{2+}–NO.

The PAS and CC domains of sGC have been studied less by comparison, though their function in sGC signal transduction is likely important. PAS domains are versatile as they can play varied roles in different proteins, such as quaternary structure organization, cofactor binding, and signal transduction.[37–39] Indeed, in sGC the PAS domains are implicated in dimerization,[40] gas binding,[41] and Hsp90-mediated heme insertion.[42] Additionally, the crystal structure of the *Manduca sexta* sGC α_1 PAS domain was recently solved, revealing an extra beta strand and a flexible helix that could be important for transducing the NO signal;[43] however, the exact mechanism of signal transduction is still not fully understood. The first crystal structure of the β CC domain in sGC exhibited the coils oriented in an anti-parallel orientation.[44] More recently, cross-linking mass spectrometry and single particle electron microscopy experiments suggest that the orientation of these helices are parallel in the native structure.[41,45] Through the use of hydrogen deuterium exchange mass spectrometry, the linker region between the PAS and the CC domains

has been shown to undergo remarkable solvent exchange dynamics with the addition of NO, suggesting that conformational changes affect this linker region in an NO-dependent manner.[46] Since the PAS/CC linker region appears to be important for transducing the NO signal, additional studies to elucidate the molecular details of this process will be very important.

The CAT domains of sGC are members of the Class III nucleotide cyclase domain family, as determined by sequence analysis and secondary structural motifs.[47–50] The CAT domains form a heterodimer: one active site and one pseudosymmetric site are formed at the interface of the α and β subunits. ATP is a mixed-type inhibitor of sGC and this pseudosymmetric site is proposed to be the site of ATP action.[17,51,52] The interested reader is directed to the review by Childers et al.[53] for more details on the structure and function of the CAT domain.

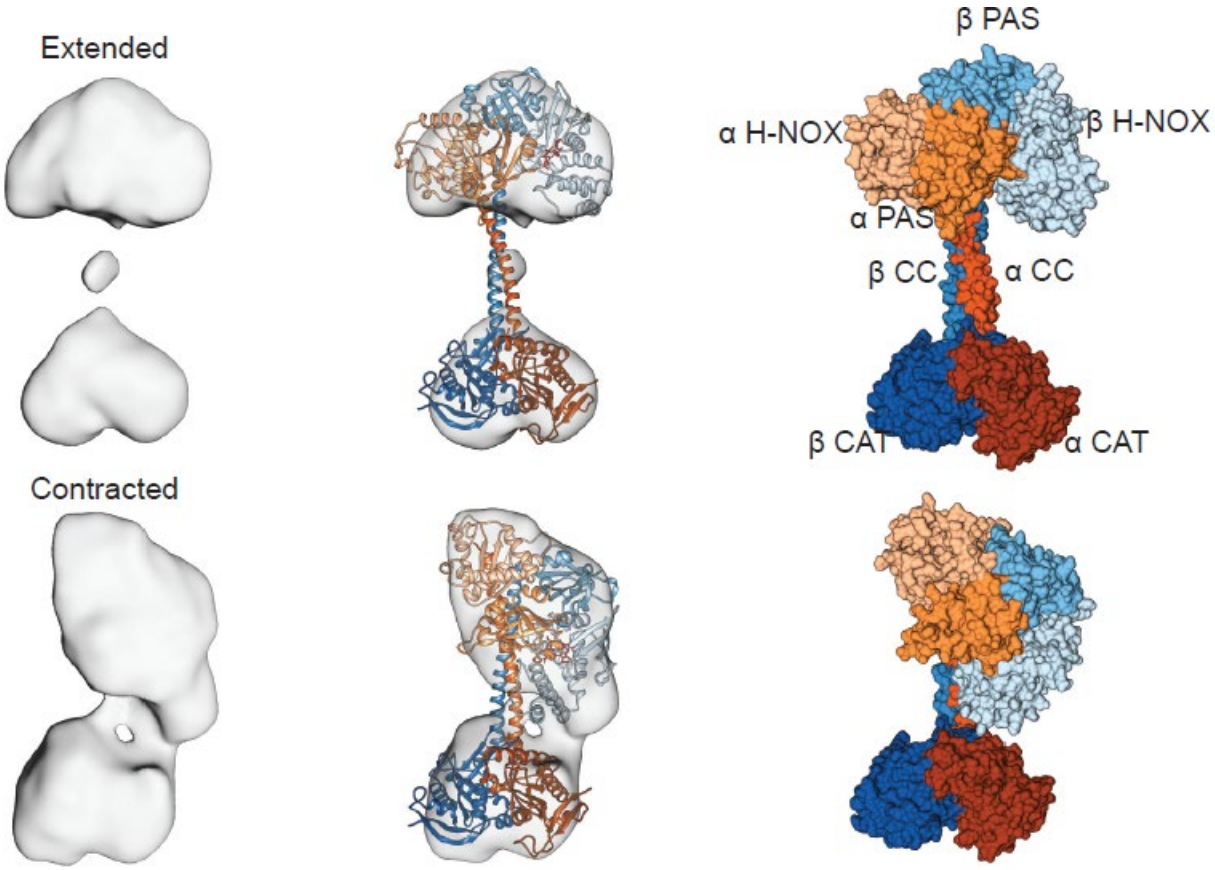


Figure 1.3 – Reconstructed full-length representation of sGC. Using available crystal structures and single-particle electron microscopy data from Ref. 40, a representation of the full-length structure of sGC was constructed. The subunits are colored as in Figure 1.1A. Beginning with two representative EM volumes (left), the following crystal structures were fit into the densities: Human α1 and β1 sGC H-NOX domain homology models were derived from the bacterial H-NOX from *Nostoc sp* (PDB ID: 2O09); Human α1 and β1 PAS domain homology models were derived from eukaryotic PAS from *C. reinhardtii* (PDB ID: 4GJ4) and aligned using the *Nostoc punctiforme* PAS dimer (PDB ID: 2P04); Human α1 and β1 CC domain homology models and the human CAT structure (PDB ID: 4NI2) were aligned to the *Mycobacterium intracellulare* adenylate cyclase structure (PDB ID: 505l), for chains C and D, and adjusted to fit within the volumes. These crystal structures have been fitted into the densities and are shown as ribbon diagrams (middle). The two extreme conformations of sGC are shown and labeled as "Extended" and "Contracted" (see Figure 1.4). The space-filling model of both conformations is shown (right).

Full-length structures of sGC have been obtained only recently after twenty-five years of work (see Chapter 2)[54], the delay likely due to conformational flexibility in the quaternary structure in conjunction with challenges of expressing sufficient quantities of protein. Previously, Campbell et al.[45] successfully used single particle electron microscopy to visualize the quaternary structure of sGC. The H-NOX and PAS domains were clustered in a heterotetramer-like lobe at one end of the particle, while the CAT domains formed a tight heterodimeric lobe on the opposing end with the CC domain forming a bridge (Figure 1.3). These studies showed that sGC is highly flexible in the presence and absence of NO, independent of the pH of the stain. Recently, Vercellino et al.[55] solved the first crystal structure containing both the CC and the CAT domains of an adenylate cyclase from *Mycobacterium intracellulare.* Because these structural motifs are conserved across many nucleotidyl cyclases, this crystal structure can be used to improve modeling of the quaternary structure of sGC, as shown in Figure 1.3.

There had been two hypotheses that predict how the quaternary structure of sGC changes upon NO activation. It was postulated that the β H-NOX domain comes in direct contact with the α CAT domain through space to inhibit CAT domain activity (Figure 1.4A). Alternatively, it was proposed that the signal transduction event occurs through the protein backbone (Figure 1.4B). The hypothesis that the β H-NOX domain inhibits the catalytic activity is based on individual H-NOX domain inhibition of catalytic activity *in trans*, and cross-linking experiments establishing contacts between these two domains.[41,56] Support for the signal transduction event to take place through the protein backbone comes from the observation of NO-independent flexibility observed by electron microscopy.[45] While it is possible that some continuum between these two extremes is physiologically relevant, important new details are now available with the recently published strtcures (Chapter 2).

Activation of Soluble Guanylate Cyclase

The molecular steps involved in sGC activation have evolved significantly over the last several years. The key studies are outlined below, but here we provide an overall summary. The activation mechanism of sGC was initially proposed to be relatively simple, where the 5c Fe^{2+}–NO heme complex generated upon NO binding led to fully active enzyme. Detailed kinetic studies on NO binding to the heme and enzyme activity has led to a more complicated mechanism, perhaps in hindsight not unexpected given the pivotal role played by sGC in human physiology. There are two aspects to NO activation of sGC. The first involves the formation the 5c ferrous nitrosyl heme complex, and the most recent evidence is consistent with a distal version of this species (5c Fe^{2+}–NO). sGC treated with 1 equivalent of NO is activated to only ~15% of maximal activity. Addition of excess NO fully activates sGC *and* increases the rate at which the 5c Fe^{2+}–NO heme complex forms from the 6c His–Fe^{2+}–NO species. All evidence points towards a second NO binding site for the action of this excess NO as discussed below.

NO binding to the H-NOX domain

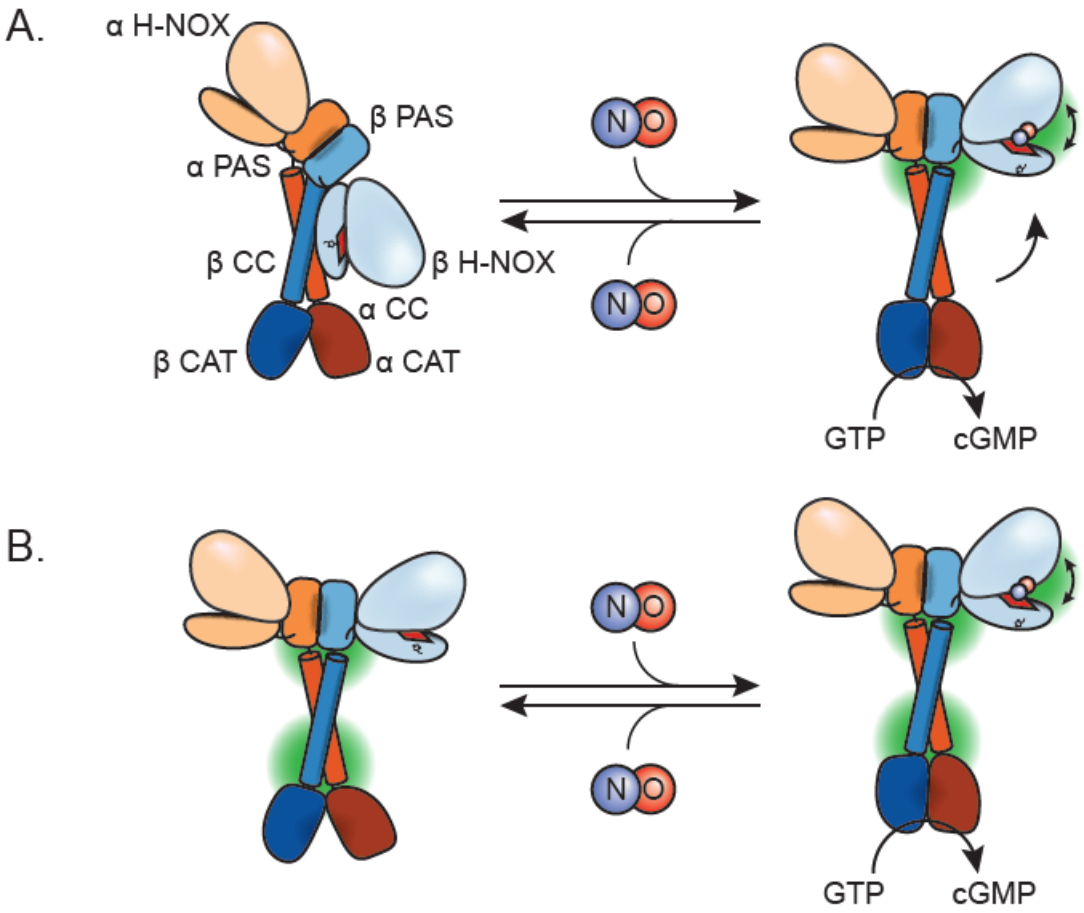


Figure 1.4 – Quaternary domain organization of sGC upon NO activation. It is predicted that some spatial reorganization of the sGC domains occurs when NO activates the enzyme. A) Self-inhibition of the CAT domain through contact of the β H-NOX domain to the α CAT domain, termed the contracted conformation. Upon NO binding, this contact would be released allowing a conformational change into the extended conformation which in full catalytic activity. B) When NO is not bound, the tertiary structure holds sGC in a non-active conformation while the quaternary structure is flexible (not depicted). Upon NO binding, structural rearrangements occur along the protein backbone, which results in full catalytic activity.

It is well established that the first step in sGC activation is NO binding to the 5c His–Fe^{2+} heme cofactor in the β1 H-NOX domain[30–32]. The resulting 6c His–Fe^{2+}–NO complex has been observed by stopped-flow UV-Vis spectroscopy and electron paramagnetic resonance (EPR) spectroscopy with samples prepared at 4 °C or lower to slow iron–histidyl bond scission.[30,57–60] The strong σ trans effect of NO ruptures the iron–histidyl bond (k_{6c} → k_{5c} = 50,000 $M^{-1}s^{-1}$), resulting in a 5c Fe^{2+}–NO heme complex, which exhibits a relatively long half-life in solution (k_{off} for NO = 0.0006 s^{-1}) compared to other species in sGC activation (discussed below).[31,61] Even so, it was only recently that Herzik et al. solved a crystal structure of the 5c ferrous nitrosyl H-NOX domain, using the H-NOX protein from the gamma proteobacterium *Shewanella oneidensis* (*So*).[34] This study also used a stable 6c His–Mn^{2+}–NO porphyrin complex to characterize possible structural changes in the H-NOX domain due to gas binding, removing the complication of iron–histidyl bond cleavage.[62] The root mean square deviation of the atoms in the 5c His–Mn^{2+} structure aligned to the 6c His–Mn^{2+}–NO structure was 0.099 Å, indicating that the driving force behind iron–histidyl bond cleavage is the σ trans effect of NO rather than any structural rearrangement

resulting from ligand binding. These results support the hypothesis that the H-NOX conformational change is dependent on iron–histidyl bond cleavage.[30,62]

Two hypotheses have been proposed for H-NOX conformational change required to initiate a gas-induced activity change in a response regulator. Ma et al. solved crystal structures of the unliganded, NO-bound, and CO-bound H-NOX protein from *Nostoc* sp., and found that in this H-NOX, both ligands formed 6c complexes. The authors proposed a heme pivoting mechanism that is dependent on gas binding for signal transduction. Herzik et al. compared the structures of *Shewanella oneidensis* with and without NO and described a relative motion of the N-terminal and C-terminal subdomains along a pair of conserved glycines, termed the "glycine hinge" (Figure 1.1B). However, the mechanism of H-NOX induced conformational change resulting in activation of sGC is currently not known and await high resolution structures of full-length sGC with and without NO bound for clarification of these possibilities (see Chapter 2).

Connecting NO binding to iron–histidyl bond cleavage

Stone et al.[57] first observed an unexpected phenomenon that indicated interaction of NO with sGC is not as simple as NO binding to the heme: the rate of conversion of 6c His–Fe^{2+}–NO to 5c Fe^{2+}–NO depends on the concentration of NO. This result was then expanded upon by Zhao et al. (Figure 1.5).[30] This surprising concentration dependence implies that a second molecule of NO must interact with the 6c His–Fe^{2+}–NO sGC enzyme to influence the rate of iron–histidyl bond dissociation. There are several proposals for this additional interaction: a second, non-heme NO binding site that could be covalent or non-covalent; nucleophilic addition of NO to the heme-bound NO; or a second NO binding to the proximal side of the heme. These proposals remain somewhat contested,[63,64] and the molecular basis for this concentration dependence has yet to be conclusively demonstrated. Given that two of the proposals are difficult to test experimentally, namely nucleophilic addition of NO to the heme-bound NO and a hydrophobic site specific for NO, in this section we review the major studies that seek to determine whether this dependence on NO concentration of histidine dissociation is due to a second NO binding event on the proximal side

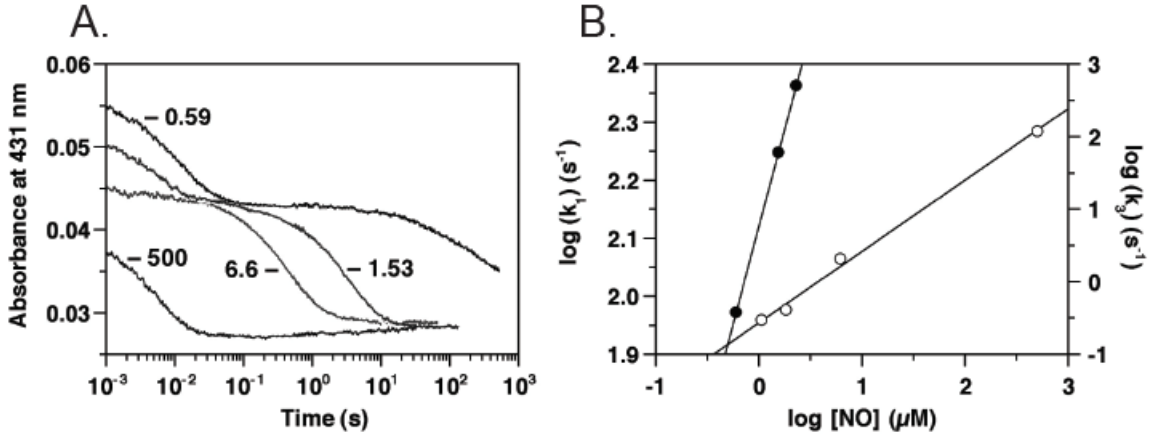


Figure 1.5 – Nitric oxide-dependent spectral transitions in sGC. Both the formation of the 6c ferrous nitrosyl complex and cleavage of the iron–histidyl bond are dependent on NO concentration. A) The absorbance at 431 nm, which corresponds to Soret maximum the 5c ferrous-histidine complex, decreases with time as sGC interacts with the indicated NO concentrations (in μM). B) Dependence of the slow and the fast rates extracted from A) on NO concentration. Adapted from ref[30].

of the heme. The next section of the review will be devoted to articles focused on a second non-heme NO binding site.

When the hypothesis that NO could bind to the proximal side of the heme was first proposed in 1999 by Zhao et al.,[30] cleavage of the iron–histidyl bond was considered to be the only step necessary to fully activate sGC. NO first must bind to the distal side of the heme forming a transient 6c His–Fe^{2+}–NO complex before iron-histidine bond cleavage leading to 5c Fe^{2+}–NO (Figure 1.6A, species (1) → (2) → (3)). Accordingly, the re-formation of the iron–histidyl bond was thought to be the only step involved in sGC deactivation. When NO binds to the distal face of the heme (5c Fe^{2+}–NO), the dissociation path is straightforward. By the principle of microscopic reversibility, NO would dissociate in the reverse path of its association (Figure 1.6A, species (3) → (2) → (1)). However, if NO binds to the proximal face of the heme (5c NO–Fe^{2+}), it is not clear how NO would dissociate, assuming the generation of a 4-coordinate porphyrin is energetically unfeasible (not depicted) (Figure 1.6B, (5) → (1)). Without excess NO (or another ligand) to bind and reform a 6c species, the dissociation of proximal NO would not be easily reversed.

Nonetheless, several studies have found NO bound in the proximal pocket of bacterial H-NOX proteins and in the H-NOX domain of sGC. In the NO-bound crystal structure of *So* H-NOX mentioned above, electron density corresponding to the NO ligand was observed on the proximal side of the heme (5c NO–Fe^{2+}). In this structure with the proximal ligation of NO, the proximal histidine was shifted 8.5 Å and the αF helix rotated 90° (Figure 1.1B). Proximal NO ligation was also observed in an earlier crystal structure of cytochrome c′ from *Alcaligenes xylosoxidans*. Although the protein structures are unrelated, this cytochrome c′, also forms a 5c NO–Fe^{2+} heme cofactor via a 6c dinitrosyl complex (6c NO–Fe^{2+}–NO)[65]. However, the crystal structure of the *So* H-NOX domain provided the first crystallographic evidence that NO can bind to the proximal side of the heme in H-NOX domains, and thus potentially in sGC. The structure of the 5c NO–Fe^{2+} H-NOX exhibited a displacement of the distal domain by 2.5 Å compared to the 5c His–Fe^{2+}structure, when the two structures were aligned by the proximal domain (Figure 1.1B). A similar displacement was also observed in prokaryotic O_2-sensing H-NOX domains, suggesting that these structural changes upon gas binding may be universal for the signal transduction of gases through H-NOX domains, including in sGC.[66]

In 2012, Martin et al.[59] found that NO could coordinate to the proximal side of the heme in full-length sGC. Rapid freeze quench (RFQ) EPR with isotopically labeled NO was used to discern which isotopes of NO bound to the heme under sequential mixing conditions. First, 5c His–Fe^{2+}sGC was mixed with ^{14}NO to generate 6c His–Fe^{2+}–^{14}NO, then ^{15}NO was added and the samples were rapidly frozen after approximately 5 seconds. Isotopically mixed 5c samples were measured by EPR, implying the formation of two different populations: ferrous nitrosyl heme with either ^{14}NO or ^{15}NO bound. Although this EPR experiment does not distinguish between NO bound to the distal or the proximal side, the experimental design was such that the NO isotope used in the second addition must initially bind to the proximal side. The authors were able to show that the ratio of ^{14}NO:^{15}NO bound to sGC was influenced by the concentration of ^{15}NO used. For example, whereas an equimolar concentration ratio of ^{14}NO:^{15}NO gave sGC populations with 43:57 (^{14}NO:^{15}NO) bound, ten-fold excess of ^{15}NO led to an sGC population of 15 : 85 (^{14}NO:^{15}NO). This concentration dependence indicates that on this timescale, NO–Fe^{2+}–NO forms, however it is not clear how many times NO is binding to the heme so NO could be proximal or distal (species (3), (4), and (5) in Figure 1.6). Under these conditions on the timescale of

milliseconds to seconds, NO is binding to the proximal side of the heme, and the authors conclude that the proximal side of the heme is the site where the second NO molecule binds.

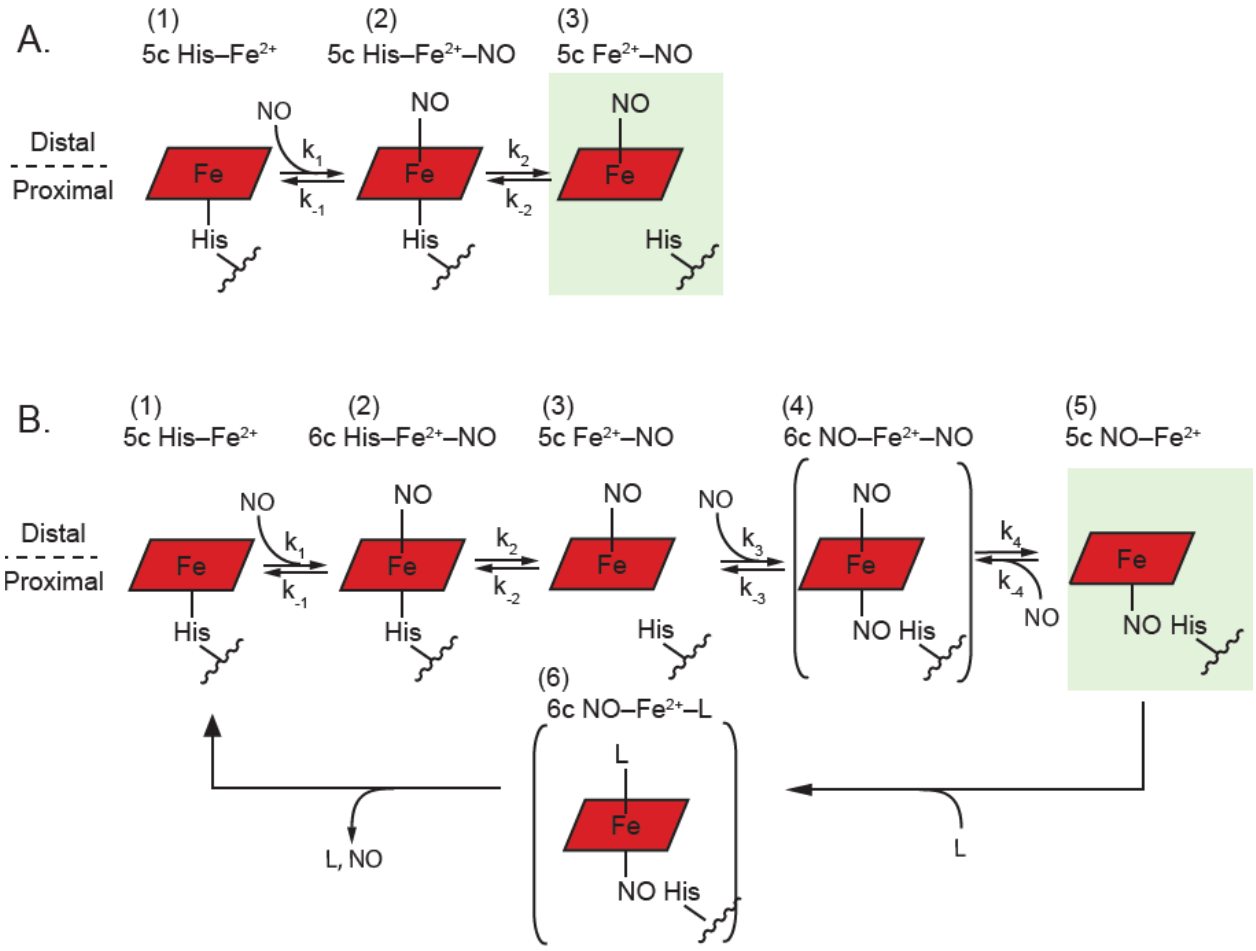


Figure 1.6 – Nitric oxide sidedness on the heme. A depiction of the molecular steps of NO binding to and dissociation from the heme in sGC. A) NO binds to the distal side of the heme with an association rate of k_1 to form species (2). Rupture of the proximal histidine iron bond with the rate of k_2 yields species (3). Note that k_2 is dependent on NO concentration, as shown in Figure 1.5. To reverse this process, the proximal histidine rebinds to the heme with a rate of k_{-2}, followed by the dissociation of NO from the distal side. B) For NO to bind on the proximal side of the heme, the distal 5c species must form first (as in A). Another molecule of NO binds to the proximal side of the heme with a rate of k_3 to form species (4), followed by the cleavage of the distal NO–Fe bond with a rate of k_4 to yield species (5). The reverse of this process requires either excess NO (to undergo the microscopic reverse), or another ligand (L) to bind to the distal side of the heme. Figure adapted from ref[67].

Conversely, Yoo et al.[31] studied the motion of the proximal histidine of sGC with NO present by using time-resolved absorption spectroscopy on timescales from nanoseconds to milliseconds, and did not see evidence of NO binding to the proximal side of the heme. Previous work by the authors established that laser flash photolysis of ferrous nitrosyl heme to generate the 4c Fe^{2+} species resulted in 97% geminately recombined ferrous nitrosyl heme (τ = 7.5 ps).[68,69] They then examined the 3% of heme that did not geminately recombine with NO; this population rebound the proximal histidine to form 5c His–Fe^{2+}(τ_0 = 70 ps). Following this 5c His–Fe^{2+} species, four specific transitions were identified spanning nanosecond to millisecond timescales (Figure

1.7). Of these, the first two spectral changes to occur were unaffected by NO concentration: 6c His–Fe^{2+}–NO formation (τ_1 = 6.5 ns), and the subsequent iron–histidyl bond scission to form 5c Fe^{2+}–NO (τ_2 = 0.66 μs). These early events were attributed to geminate recombination of NO to the heme after the laser pulse has ended, and subsequent iron–histidyl bond rupture. The spectral changes resulting from geminate recombination of NO (τ_1 and τ_2) are an outcome of the experimental setup where the initial condition of ferrous nitrosyl sGC is followed by NO dissociation from the heme and are not relevant to physiological activation of sGC. Additional processes occur on longer timescales and are dependent on NO concentration: 6c His–Fe^{2+}–NO formation (τ_3) and iron–histidyl bond scission (τ_4). Concentrations of 20 μM and 200 μM were used, bracketing the lifetimes of these processes at 50 μs ≤ τ_3 ≤ 250 μs and 10 ms ≤ τ_4 ≤ 43 ms, which are consistent with bimolecular rates measured by stopped-flow UV-Visible spectroscopy.[30] The authors conclude that these four observed kinetic transitions can be described by a model where NO only binds to the distal side of the heme, at least within a window of 200 ms following initial NO binding.

Importantly, Yoo et al.[31] observed that the slow iron–histidyl bond rupture (τ_4) is NO-dependent, consistent with results from Zhao et al.[30] The transition that is dependent on NO concentration is strictly 6c His–Fe^{2+}–NO to 5c Fe^{2+}–NO. Since no spectral signature of the 6c NO–Fe^{2+}–NO was observed (or any higher Fe^{2+} coordination species as has been postulated), the hypothesis that the NO dependence of the iron–histidyl bond scission is caused by NO binding to the proximal side of the heme is not supported by this data.[59] It is noteworthy that the lifetime of the fast iron-histidyl bond scission (τ_2) is independent of NO concentration and is faster than the predicted timescale for protein conformational changes (1 – 50 μs), suggesting there is an additional, unobserved change in sGC that has different effects on the fast and slow iron–histidyl bond dissociation rates. However, for the reasons outlined above, this change cannot be a result of proximal NO binding. It has also been shown that treating sGC with excess NO then rapidly quenching the reaction activates sGC on sub-second time scales, as fast as 200 ms.[30] When evaluated in the context of the time-resolved data, this suggests that bimolecular NO binding and iron–histidyl bond cleavage occurs prior to sGC activation, but before proximal NO formation.

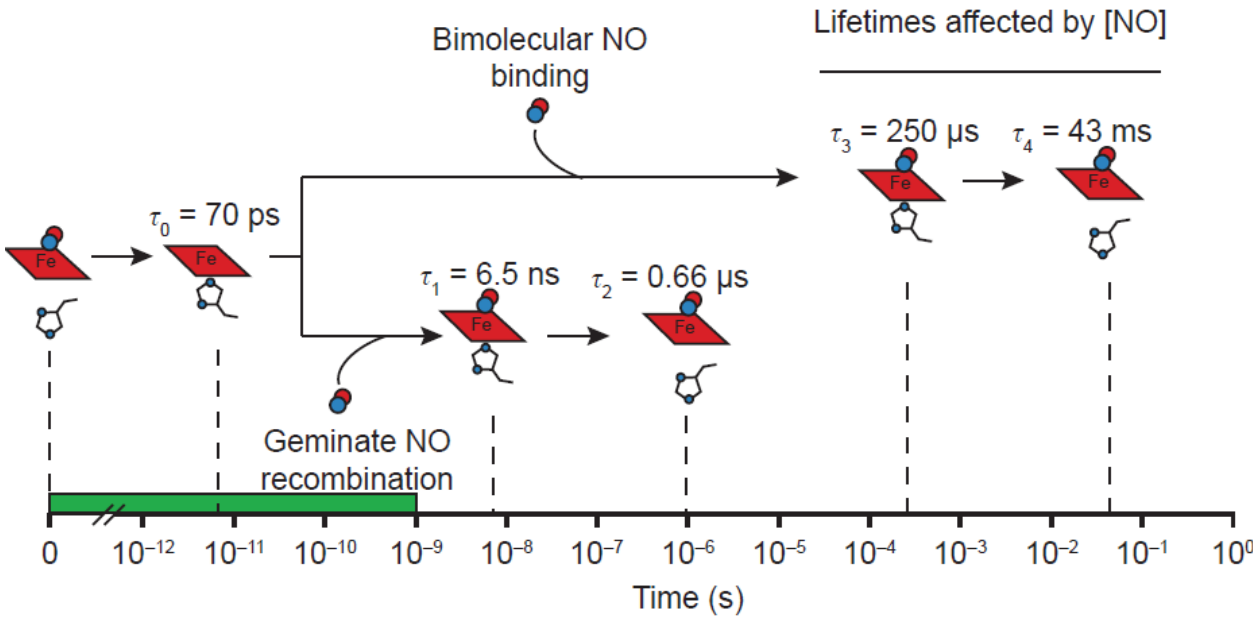


Figure 1.7 – Kinetics of sGC transitions in the heme pocket. The fate of the 5c iron-histidine species formed after laser photolysis of NO is depicted. A 6 ns 532nm laser pulse, indicated by the green bar, dissociates NO. The proximal histidine rebinds to the heme with a time constant of τ_0, measured in a separate experiment. There are two routes through which the 5c His–Fe^{2+} species can proceed. In the route occurring at faster timescales (bottom), geminate recombination of NO occurs with a lifetime of 6.5 ns, followed by cleavage of the iron–histidyl bond with a life time of 0.66 µs. τ_1 and τ_2 do not depend on NO concentration. In the route occurring at longer timescales (top), NO from the solution diffuses into the heme pocket in a bimolecular process and binds to the heme with lifetimes of 50 µs $\leq \tau_3 \leq$ 250 µs, followed by cleavage of the iron–histidyl bond with lifetimes of 10 ms $\leq \tau_4 \leq$ 43 ms (NO concentration between 20 µM and 200 µM). Both τ_3 and τ_4 are dependent on the NO concentration, consistent with the stopped-flow spectroscopy data in Figure 1.5. Figure adapted from ref[81].

On timescales shorter than a second, it appears that sGC does not form a proximal NO complex. However, such a complex may form under steady-state conditions. One result of this hypothesis would be that the 5c NO–Fe^{2+} species would require excess NO to access 5c Fe^{2+}–NO via a 6-coordinate intermediate (Figure 1.6B, species (5) → (3)), and thus would be predicted to dissociate more rapidly with excess NO than with substoichiometric NO. To evaluate whether the proximal NO complex forms under steady-state conditions, Guo et al.[67] measured the NO off-rate for a bacterial H-NOX from *S. oneidensis* under steady-state conditions with substoichiometric or excess NO. However, the observed off-rates were the same within experimental error and are inconsistent with a model where the primary species in solution is 5c NO–Fe^{2+}. Double electron-electron resonance electron paramagnetic resonance (DEER-EPR) spectroscopy was then used to measure the distance from NO to a spin label bound to a surface cysteine residue located 23 Å from the distal side of the heme and 28 Å from the proximal side of the heme. This experiment enabled the differentiation of NO bound to the proximal and distal sides of the heme. While some proximal NO is observed in the presence of excess NO, the major species was in all cases 5c Fe^{2+}–NO. In fact, when treated with excess NO and then that excess was removed, NO in that case was found only in the distal position.

Given the disparate conclusions from these studies, is it possible to reconcile these differences into a unified model for sGC activation? One hypothesis is that the proximal NO-bound species is more rigid than the distal analog, which enabled the selective crystallization of proximal-bound NO H-NOX. Both the RFQ-EPR and DEER-EPR studies indicate that the proximal NO species can form with sufficient time and NO concentration, however, the primary species at equilibrium is the distal NO complex. If the proximal NO complex were to form, it would occur after the iron–histidyl bond has been broken and sGC has already been activated, given the timescale of these processes. Taken together, it appears that NO binding to the proximal side of the heme is likely not the reason for the NO dependence of iron–histidyl bond scission which, in turn, is necessary but not sufficient for physiological activation of sGC.

NO binding to the second site of sGC for full physiological activation

Alternative hypotheses exist to explain the NO dependence of the iron–histidyl bond cleavage. Specifically, several groups have postulated that a second, non-heme NO binding site exists elsewhere on sGC. Here, the biochemical evidence for this secondary binding site will be discussed.

The first biochemical evidence for a second NO binding site was characterized by Bellamy et al. in 2002.[70] The authors measured the activation of sGC using constant or clamped NO concentrations (where the addition of a NO donor is in equilibrium with a NO scavenger) and found that the Hill coefficient of sGC to be 2.1. This suggests that there is cooperativity in the binding of more than one NO molecule to sGC, and the molecular steps in full activation of sGC involve more than NO binding to the heme. In 2004, Russwurm et al.[71] measured both the UV-Vis spectrum and the activity of sGC in the same sample. Upon the addition of excess NO, the UV-Visible spectrum remained unchanged, but the activity increased dramatically compared to substoichiometric amounts of NO. The authors found they could generate a low activity state of sGC by either adding a sub-stoichiometric amount of NO relative to sGC, or by removing excess, non-heme bound NO by a buffer exchange. These results suggested that sGC exists in two different activity states, depending on the concentration of NO, without any changes to the 5c ferrous-nitrosyl heme. The proposed second site has a dissociation constant much higher than that of the heme cofactor ($K_d < 1.2 \times 10^{-12}$ M).[30] The substiochiometric, low-activity state will be henceforth referred to as the 1-NO state, indicating that one NO is bound to sGC at the heme and the enzyme is in a low activity state. The occupancy of the heme by NO under these conditions is 100%.

In 2005, Cary et al.[17] compared enzymatic activity to the dissociation rate of NO in sGC. The authors found that sGC deactivates over 150 times faster than NO dissociates from the heme with purified enzyme, indicating that NO dissociation does not directly cause sGC deactivation. These results suggest that a binary model of sGC activation and deactivation is likely incomplete. Corroborating Russwurm's data, the authors isolated 1-NO sGC with low activity by using two different traps that bind excess, non-heme bound NO, further indicating that breaking the His-Fe^{2+} bond is not sufficient to fully activate sGC. Additionally, the authors found that YC-1, a known stimulator of sGC activity, fully activates the 1-NO state and has only a small effect on the activity of sGC when excess NO is present. Subsequent deactivation of the 1-NO sGC with YC-1 present occurs slowly. This result is critical to discussing the activation of sGC in cells, discussed below. The model proposed in this paper is a three-activity model of sGC, where sGC has basal activity

without NO, ~15% activity with one equivalent of NO, and 100% activity when excess NO is present.

In a slightly different approach, Derbyshire et al.[72] used butyl isocyanate (BIC) as a ferrous distal heme ligand (6c His–Fe^{2+}–BIC) to block diatomic gases from binding to the heme. With BIC as the distal ligand, it was found that the addition of excess NO resulted in a slight shift in the Soret band from 429 nm to 432 nm. This result indicates that excess NO affected the heme environment without replacing the BIC ligand, as ferrous nitrosyl heme would exhibit a Soret band at 399 nm. BIC-bound sGC was activated 5-fold over basal activity in the absence of NO, compared to BIC-bound sGC with excess NO was activated 89-fold over basal. Addition of dithionite, which reacts with all non-heme bound NO, eliminated this additional activity. Together, these findings strongly suggest that there is a second, non-heme NO binding site whose occupancy is required for full sGC activation.

Fernhoff et al.[73] hypothesized that the second NO molecule interacts with a cysteine residue on sGC. In this study, the authors used *S*-methyl methanethiosulfonate (MMTS), a small molecule that reversibly reacts with thiols to form an *S*-methyl-cysteine disulfide bond. When the cysteines of sGC formed a disulfide bond with MMTS, the activity was ~15% of maximal activity with excess NO (Figure 1.8A) and is nearly identical to the activity of the 1-NO state (Figure 1.8D). Under these conditions, all heme was found to be in the ferrous nitrosyl state. Maximal activity was restored upon the addition of excess thiol to remove the MMTS label, thus returning the cysteine disulfides to the native thiol (R-SH) (Figure 1.8C). Excess MMTS did not affect the basal activity of sGC (Figure 1.8B).

The interaction of thiols and NO is not a new idea. *S*-nitrosation is a known chemical modification of proteins that has been studied for nearly two decades. However, formation of *S*-nitrosothiols (RSNO) requires a one-electron oxidation of NO and is thus dependent on the presence of an oxidant. Conversely, sGC activation does not dependent on oxidation.[73] Additionally, the half-life of an *S*-nitrosothiol is on the order of hours, which is far longer than the timescale of sGC deactivation which occurs in seconds. Therefore, the authors proposed that NO interacts with an sGC cysteine as a transient *S*-nitrosylthiol radical anion ($RSNO^{\bullet-}$). While $RSNO^{\bullet-}$ and RNSO differ by only one electron, that single electron dramatically changes the properties of the complex. The half-life of an *S*-nitrosylthiol radical anion could be stabilized by adjacent residues via electrostatic interactions. Additionally, the nucleophilic addition of a thiolate to NO to generate $RSNO^{\bullet-}$ is freely reversible. Such an interaction is commensurate with the concentration of free NO in cells and provides a rapid physiologically-relevant deactivation mechanism of sGC. Since the critical cysteine(s) involved in this transient interaction have yet to be identified, the NO–cysteine interaction remains a hypothesis, albeit with compelling evidence to support it.

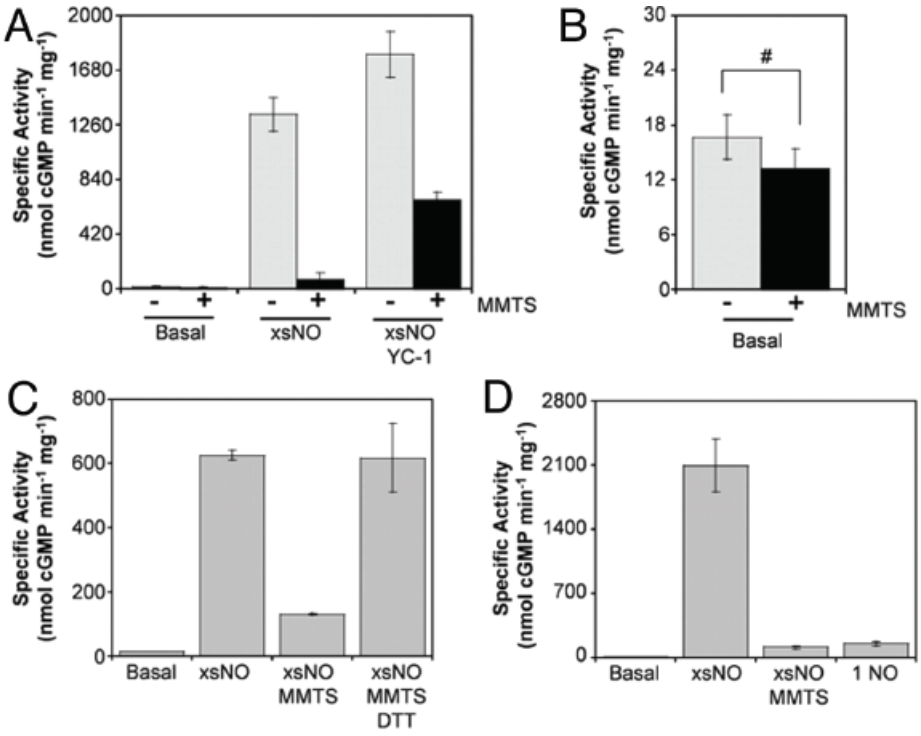


Figure 1.8 – Blocking of sGC cysteines prevents full enzyme activation. A) Experimental data showing that MMTS-treated sGC will not fully activate with excess NO but will activate to a certain extent with NO and a stimulator. B) MMTS does not affect the basal activity level of sGC. C) Treatment of the MMTS-labeled enzyme with DTT reverses the inhibition. D) The activity of the MMTS-treated sGC is comparable to the activity of the 1-NO state. Figure reprinted from ref[73].

An NO–cysteine interaction could explain the NO-dependence of τ_4, but not τ_2, as observed by Yoo et al.[31] Those authors described a structural change in sGC that occurs on a timescale between τ_2 and τ_3 in order to account for the vastly different time scales of iron–histidyl bond cleavage. However, no explanation why τ_4, but not τ_2, depends on NO concentration was provided. An interaction between NO and a cysteine is likely to occur on this timescale and induce the observed structural transition. This conformational change could then influence the cleavage of the iron–histidyl bond, thereby accounting for the NO-dependence of τ_4.

Cary et al.[17] showed that small molecule sGC stimulators synergize with the 1-NO state of the enzyme. However, discovering precisely how NO and stimulators activate sGC to maximal activity is of great interest and directly relevant to the therapeutic action of these molecules. Despite a number of studies that seek to identify or narrow down the site of stimulator binding,[74–82] the mechanism by which sGC activity is increased is still unknown. A recent study by Wales et al.[83] localizes the stimulator binding site between the N- and C- terminal subdomains of the β1 H-NOX domain, clarifying the specificity of these stimulators to sGCs. Further studies focused on the interaction between stimulators and sGC residues will parse out the necessary steps for stimulator activation and their synergy with NO.

sGC activation in cells

sGC activation should occur on similar timescales with the same ligand dependence, whether measured with purified enzyme or in cells. If this is not the case, it could signify that

additional cofactors or binding partners are missing with studies conducted with purified enzyme. Measuring sGC activity directly in cells is challenging experimentally, and yet significant progress has been made toward this goal. Bellamy et al.[84] constructed an apparatus to reliably quench rat cerebellar cell suspensions on millisecond timescales. No lag in the synthesis of cGMP was detected, even at 20 ms after addition of NO, indicating that sGC is activated on sub-millisecond timescales at cellular concentrations of sGC. This is in agreement with the timescale of sGC activation as measured with purified enzyme.[30]

Concerning the ligation state of sGC, there is currently no accurate method to measure the ligation state of heme cofactors *in vivo*. However, the assumption that sGC in cells has NO bound to the β H-NOX domain is based on two different sets of experiments. The first is that free NO has been measured by a variety of techniques, including fluorescent probes, NO biosensors, and several other methods. These studies indicate that the physiological signaling concentrations of NO are between 100 pM to 5 nM.[85–89] This is 100 to 1,000-fold higher than the dissociation constant for NO bound to the H-NOX domain of sGC ($K_d < 1.2 \times 10^{-12}$ M).[30] Based on these values, it seems likely that the resting ligation state of sGC in cells is 5c Fe^{2+}–NO. The second set of experiments examined the effect of adding a small molecule sGC stimulator to human umbilical vein endothelial cells.[73] Without the addition of excess NO, YC-1 was able to stimulate sGC activity. As work with purified enzyme shows YC-1 has very little effect on 5c His–Fe^{2+} sGC, this implies there is a ligand` bound to the heme cofactor of sGC in cells.

As discussed, measuring the ligation state of sGC in cells has been a formidable challenge. From work with purified enzyme, NO dissociates slowly from the heme, even in the presence of ATP and GTP.[17] However, either with purified enzyme or experiments done in cells, sGC activity quickly decreases once NO is removed.[17,71,87,89] Roy et al. measured the EC_{50} of sGC in rat platelets to be 11 ± 1 nM, with a Hill coefficient of 1.2 ± 0.1, and the authors concluded that sGC possesses one binding site for NO in this concentration range (approximately 1 nM to 200 nM NO).[87] One interpretation from these data is that the one binding site measured was for the heme. However given that the K_d of the H-NOX domain for NO is four orders of magnitude below the EC_{50}[30] and the Hill coefficient for NO activating purified sGC is 2.1,[70] another interpretation of these data is that sGC in cells already has NO at the heme and the one binding site measured is the second site cysteine interaction.

Using the recently developed fluorescent cGMP biosensor δ-FlincG, Batchelor et al.[88] measured sGC activation at very low concentrations of NO, in the picomolar range. They developed a model that accounts for both sGC activation and the phosphodiesterase activity that degrades the cGMP signal. An increase in activity at such a low NO concentration is unusual, but the authors argue that the efficiency of signal transduction (defined as the turnover of enzyme per molecule of NO) is highest under these conditions. These experiments demonstrate that sGC can respond to exquisitely low amounts of NO, much lower than previously thought. The EC_{50} for NO in these studies depended on the relative concentrations of sGC and phosphodiesterase, however it is likely between 1-10 nM, in agreement with other studies done in platelets and astrocytes.[90,91] Thus, studies conducted in cells are consistent with work with purified enzyme and demonstrate that sGC activates on sub-second timescales, the ligation state of the H-NOX is likely 5c Fe^{2+}–NO, and sGC can respond to low amounts of NO.

Activation of Atypical Soluble Guanylate Cyclases

In addition to the NO signaling pathway elucidated in mammals, there are homologous signaling pathways in lower animals that use NO and O_2 as primary signaling molecules. For example, specific invertebrates contain either NO-sensitive sGCs, O_2-sensitive sGCs, or some species carry both NO and O_2 sensitive sGCs. The fall army worm (*Manduca sexta*) expresses NOS and sGC within neuronal cells where the signaling pathway is thought to play critical roles in the development of the olfactory system. Soil-dwelling nemotodes including *Caenorhabditis elegans* have been shown to have O_2-sensing sGCs that respond to environmental O_2 concentration to adjust social and eating behaviors.[92] *Drosophila melanogaster* possesses both NO-sensing and O_2-sensing sGCs, however distinct spatial expression patterns segregate these sensors to distinct locations in larvae.[93] There are also some reports of carbon monoxide (CO) acting as the physiological ligand for sGCs. Iron-limited cultures of *Chlamydomonas reinhardtii* produce limited amounts of chlorophyll. However, production of chlorophyll recovers upon addition of CO, indicating some sensing mechanism for this gas is present in the organism (see Chapter 4).

Deactivation of Soluble Guanylate Cyclase

Physiological deactivation of sGC

GC deactivation is rapid following the removal of free, excess NO, both in cells and with purified enzyme.[84,94] To understand deactivation, the physiological concentrations of NO and sGC in cells must be considered. A recent estimate of sGC concentration in cells is 2 μM, deduced from cGMP formation in platelets using the kinetic parameters of isolated sGC.[88] Additionally, by deleting one isoform of the α-subunit of heterodimeric sGC and measuring the vascular response in both control and knockout mice, one study found that fewer copies of sGC can still elicit a full physiological response.[95] Thus, there is general agreement that sGC is present in excess relative to NO in cells.

As mentioned above, the ligation state of sGC in cells is assumed to be 5c Fe^{2+}–NO. This hypothesis stands in apparent contrast to the measured intracellular abundances of sGC (~2 μM) and NO (~1 nM); however, it should be noted that measurements of NO in cells is likely an underestimation as these data come from measurements made on free NO and thus sGC-bound NO is not accounted for in these estimations. Thus, since the measured concentration of free NO is higher than the dissociation constant of NO from the sGC heme ($K_d < 1.2 \times 10^{-12}$ M)[30], we assume that all sGC in cells will have NO bound. Additionally, the half-life of most human proteins is on the order of hours, so even if not every sGC heterodimer is bound by NO immediately post translation, the majority of the population will exist long enough to bind to NO.[96] Consequently, all sGC in cells would be in the low-activity, 1-NO state (Figure 1.9, middle). When the local concentration of free NO increases to nanomolar concentrations, NO will interact with the second binding site, and a small fraction of sGC will be converted to the fully active state. The estimation of nanomolar concentrations as the apparent K_d of the second site is deduced from the concentration of NO necessary to fully activate both purified sGC[57] and sGC in cells[87]. In this model, deactivation of this fraction is straightforward: as the concentration of NO decreases below

1 nM within a cell, the dissociation constant for the second site will favor NO release, returning sGC to the 1-NO state.

sGC stimulators were initially characterized as NO-independent, heme-dependent modulators of sGC activity. Indeed, in cell studies, stimulators increase sGC activity in the absence of excess NO. However, with purified enzyme, ligand binding to the heme is required to increase sGC activity, although the ligand can be either NO or CO. These differences can be reconciled by returning to the three-activity model. In cells, where the ligation state of the heme is assumed to be 5c Fe^{2+}–NO, the exogenous stimulator alone can render sGC fully active without the addition of additional NO. This model was further supported by cell studies where human umbilical vein endothelial cells (HUVECS) were pre-treated for two hours with a nitric oxide synthase inhibitor N^{ω}-nitro-L-arginine methyl ester, essentially removing NO from the cells. Upon addition of YC-1, it was found that sGC activity in cells, as measured by cGMP production, was severely decreased,[73] suggesting that a heme ligand is also necessary to activate sGC with YC-1. Deactivation proceeds once the stimulator concentration decreases below stimulator's K_d for sGC.

Pathophysiological deactivation of sGC

Low sGC activity has been implicated in several pathologies, including cardiopulmonary disease, hypertension, asthma, and neurodegeneration.[13,97,98] In particular, several studies have shown that prolonged oxidative stress deactivates sGC and renders it less responsive to NO. Currently, two different mechanisms have been hypothesized to explain these effects. First, direct oxidation of cysteine thiols to *S*-nitrosothiols could decrease sensitivity to NO by obscuring the putative secondary binding site, thereby preventing a necessary conformational change or decreasing substrate binding.[99] Second, oxidation of the ferrous heme could yield ferric- and subsequently apo-sGC.[100] The discussion here will focus on the molecular details by which sGC deactivates in these pathologies.

Sayed et al.[101] evaluated the extent of cysteine *S*-nitrosation of sGC in cells when exposed to excess NO. The authors treated smooth muscle cells,cells from rat aorta, and HUVECs to varying forms of NO, including *S*-nitrosocysteine, vascular endothelial growth factor, and acetylcholine. The authors found that sGC was *S*-nitrosated and desensitized to NO. While it is known that sGC requires reduced cysteines to function, this study was the first to show that *S*-nitrosation directly leads to desensitized sGC. In a follow-up study, neonatal cardiomyocytes were exposed to NO to carry out the *S*-nitrosations in a biologically-relevant environment.[102] After examining the enriched peptides by mass spectrometry, nine new sites of *S*-nitrosation on sGC were identified. While the majority of these sites were localized to the CAT domain, two were identified in the active site pocket, suggesting that decreased substrate binding may cause desensitization.

Heme oxidation can also affect sGC activity. A recent paper by Surmeli et al.[103] provided the first biochemical data showing that ferric sGC loses heme more readily than ferrous sGC, confirming a longstanding hypothesis.[104] Apo-sGC is then quickly degraded by cells.[105] One study has sought to unify these two modes of deactivation, where ferric-nitrosyl heme of sGC may undergo reductive elimination to generate a ferrous heme and an *S*-nitrosated cysteine.[106] The authors show that exogenous thiols prevent this *S*-nitrosation event from occurring with ferric-

nitrosyl sGC, restoring sGC sensitivity to NO. However additional studies are necessary to confirm this hypothesis.

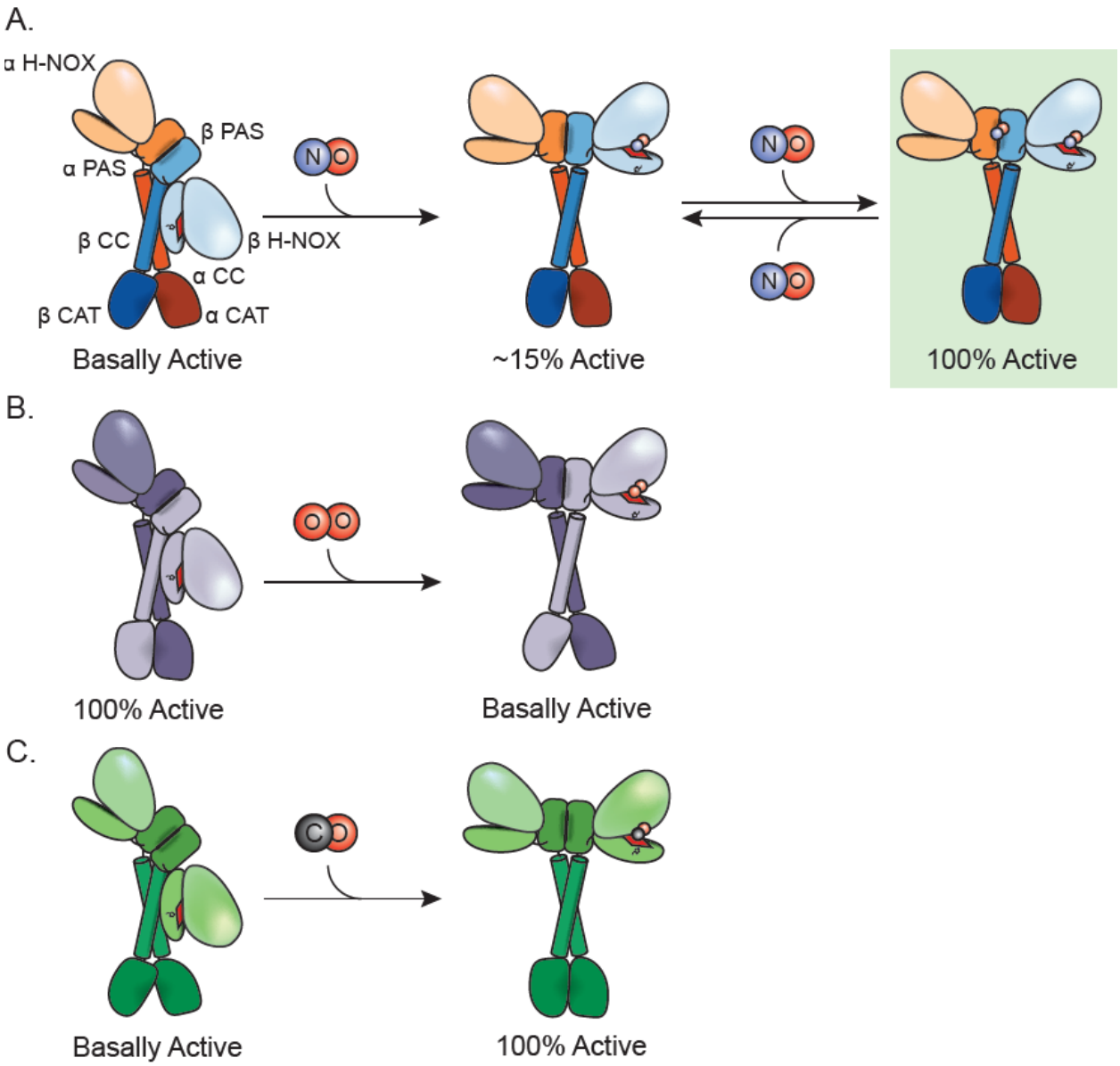


Figure 1.9 – sGC activation and deactivation *in vivo*. A) A model that accounts for all available data of physiological sGC activation and deactivation is depicted. For ease of representation, the activity of basal sGC is depicted in the conformation in which the β H-NOX contacts the α sGC CAT domain. The first molecule of NO that sGC encounters will bind to the heme of the H-NOX domain with picomolar affinity and remains bound. Thus, the vast majority of sGC in cells will be in the ~15% active state (1-NO). With an increase in the NO cellular concentration, NO will interact with the second site, which has nanomolar affinity. The physiologically relevant sGC states are boxed. Although the location of the second site has not yet been located yet, it is depicted here in the PAS domains. The modulation between these two states accounts for the downstream effects of NO in cells, as well as the rapid activation and deactivation of sGC observed both with purified enzyme, in cells, and in tissues. B) Atypical sGCs that sense oxygen are maximally active with no ligand bound, which are then inhibited with oxygen bound. C) A new class of sGC homologues have been discovered that are maximally activated with CO as the ligand.

Conclusion

The eukaryotic nitric oxide receptor sGC plays a pivotal role in eukaryotic gas signaling and is essential for many physiological processes that affect human health. The biochemical activation and deactivation of sGC has been challenging to understand, partly because of the difficulty in generating large quantities of homogeneous protein, but more so due to the complexity of the activation mechanism of the enzyme. To date, the best supported model of sGC activation and deactivation is outlined in Figure 1.9. In this three-activity model, the first molecule of NO binds on the distal side of the heme in the H-NOX domain, resulting in a partially activated enzyme. As the concentration of NO increases to nanomolar levels, NO interacts with the second site on sGC to fully activate the enzyme. As this concentration decreases, NO dissociates from the second site, and returns the enzyme to partial activity.

Moving forward, biochemical investigations centered on sGC should be validated on the full-length enzyme whenever possible to ensure catalytic competence. While truncations of sGC are easier to obtain, the full-length heterodimer is the only known form to exhibit the intricate activation sequence. Advances in structural biology techniques such as cryo-electron microscopy in the past few years have provided a high-resolution quaternary structure of sGC and will aid in answering some of the most difficult outstanding biochemical questions. Still, great care must be exercised in designing experiments to gain further understanding of this complex signaling enzyme.

Thesis Research

The focus of this dissertation centers on the activation mechanism of mammalian, insect, and algal homologues of soluble guanylate cyclase. While exclusively studying the mammalian homologue has been the norm in the first few decades, there are certain drawbacks to this narrow focus. Difficulties with protein expression and protein stability in certain biophysical experiments has increased interest in homologues of sGC. Thus, an examination of a broader panel of sGC homologues has uncovered important aspects of sGC activation by NO and YC-1, as well as a new class of sGC homologues that are more activate by carbon monoxide (CO).

Chapter 1 discusses and provides further background for the mammalian nitric oxide receptor soluble guanylate cyclase, and the activation and deactivation of the enzyme as related to a physiological setting. A detailed discussion outlines the structural aspects of sGC critical for its activation and the molecular steps involved in the activation of sGC. Emphasis is placed on NO binding to the heme and potential for the second binding site of NO. Studies are placed in context of the activation mechanism and studies of sGC in cells are highlighted. Finally, critical experiments are outlined that represent the next period of research into the mammalian NO receptor.

Chapter 2 describes the full-length structure of an insect NO-sensitive soluble guanylate cyclase from *Manduca sexta* (*Ms*). The biochemical properties of *Ms* sGC are confirmed to be similar to that of mammalian sGC, namely the binding of NO and CO, the ability for the enzyme to be stimulated by excess NO, and the ability for YC-1 to activate the 1-NO state. Cryo-EM was used to solve a 5.1 Å map of the unliganded state and a 5.8 Å map of the excess NO with YC-1 state. The unliganded state of sGC has bent coiled coil domains and the catalytic domain in a closed conformation that does not allow the substrate to bind. The excess NO plus YC-1 state of sGC has straight coiled coils, an open catalytic domain, and YC-1 bound in a central pocket formed by the β H-NOX and both CC domains. The extension of the activated state was also observed in solution by Small Angle X-Ray Scattering (SAXS). This work delineates the endpoints of the allosteric activation of sGC by NO and YC-1 and allows for detailed biochemical experiments to now be guided by structure, a long-standing goal in the field.

Chapter 3 outlines the progress made towards an understanding of the second NO binding site of sGC. Excess NO is required to generate the maximally active form of the enzyme, however certain treatments of sGC with bioconjugational agents prevent excess NO from generating this form of the enzyme. A wider range of bioconjugational agents has been explored, yet only cysteine modifications disrupt the NO-stimulated activity of sGC. The reactivities of different cysteine reactive agents has been explored to determine the requirements to disrupt activity. Structural guidance from the structures discussed in Chapter 2 and future directions will be proposed to more accurately understand the mechanism of inhibition of sGC by these bioconjugational agents.

Chapter 4 describes the characterization of a new family of sGCs that sense carbon monoxide (CO). Organisms are required to sense their external environment in order to adapt to low nutrient conditions. In low iron conditions, *Chlamydomonas reinhardtii* has been described to use CO as a signaling molecule to increase its chlorophyll production. Four of the six sGC coding genes were expressed in *E. coli* and examined biochemically. While only two of the four genes

had significant heme incorporation, it appears that the homodimers were binding one heme, despite having two binding sites. One gene, Cyg11, was activated to a larger extent by CO than by NO and had the smallest K_d for CO of any full-length sGC. This study classified the biochemistry the first CO sensor of the sGC family, opening up the potential for CO to be a signaling molecule similar to NO.

Chapter 5 outlines experiments related to Chapter 4 but did not merit publication including the $k_{on,NO}$ of Cyg11, the growth and culturing of *Chlamydomonas reinhardtii*, and the quantification of *cyg* genes transcribed. Additionally, future directions for the biochemical studies of soluble guanylate cyclases are discussed.

References

(1) Horst, B. G.; Marletta, M. A. Physiological Activation and Deactivation of Soluble Guanylate Cyclase. *Nitric Oxide* **2018**, *77*, 65–74.

(2) Furchgott, R. F.; Zawadzki, J. V. The Obligatory Role of Endothelial Cells in the Relaxation of Arterial Smooth Muscle by Acetylcholine. *Nature* **1980**, *288*, 373–376.

(3) Palmer, R. M.; Ferrige, A. G.; Moncada, S. Nitric Oxide Release Accounts for the Biological Activity of Endothelium-Derived Relaxing Factor. *Nature* **1987**, *327* (6122), 524–526.

(4) Ignarro, L. J.; Byrns, R. E.; Buga, G. M.; Wood, K. S. Endothelium-Derived Relaxing Factor from Pulmonary Artery and Vein Possesses Pharmacologic and Chemical Properties Identical to Those of Nitric Oxide Radical. *Circ. Res.* **1987**, *61* (6), 866–879.

(5) Marletta, M. A.; Yoon, P. S.; Iyengar, R.; Leaf, C. D.; Wishnok, J. S. Macrophage Oxidation of L-Arginine to Nitrite and Nitrate: Nitric Oxide Is an Intermediate. *Biochemistry* **1988**, *27* (24), 8706–8711.

(6) Knowles, R. G.; Palacios, M.; Palmer, R. M.; Moncada, S. Formation of Nitric Oxide from L-Arginine in the Central Nervous System: A Transduction Mechanism for Stimulation of the Soluble Guanylate Cyclase. *Proc. Natl. Acad. Sci. U. S. A.* **1989**, *86* (13), 5159–5162.

(7) Moncada, S.; Palmer, R. M.; Higgs, E. A. Nitric Oxide: Physiology, Pathophysiology, and Pharmacology. *Pharmacol. Rev.* **1991**, *43* (2), 109–142.

(8) Steinert, J. R.; Chernova, T.; Forsythe, I. D. Nitric Oxide Signaling in Brain Function, Dysfunction, and Dementia. *Neurosci. 16* (4), 435–452.

(9) Gibson, N. J.; Nighorn, A. Expression of Nitric Oxide Synthase and Soluble Guanylyl Cyclase in the Developing Olfactory System of Manduca Sexta. *J. Comp. Neurol.* **2000**, *422* (2), 191–205.

(10) Hervé, D.; Philippi, A.; Belbouab, R.; Zerah, M.; Chabrier, S.; Collardeau-Frachon, S.; Bergametti, F.; Essongue, A.; Berrou, E.; Krivosic, V.; et al. Loss of A1β1 Soluble Guanylate Cyclase, the Major Nitric Oxide Receptor, Leads to Moyamoya and Achalasia. *Am. J. Hum. Genet.* **2014**, *94* (3), 385–394.

(11) Wen, H.-C.; Chuu, C.-P.; Chen, C.-Y.; Shiah, S.-G.; Kung, H.-J.; King, K.-L.; Su, L.-C.; Chang, S.-C.; Chang, C.-H. Elevation of Soluble Guanylate Cyclase Suppresses Proliferation and Survival of Human Breast Cancer Cells. *PLoS One* **2015**, *10* (4), e0125518.

(12) Arnold, W. P.; Mittal, C. K.; Katsuki, S.; Murad, F. Nitric Oxide Activates Guanylate Cyclase and Increases Guanosine 3':5'-Cyclic Monophosphate Levels in Various Tissue Preparations. *Proc. Natl. Acad. Sci. U. S. A.* **1977**, *74* (8), 3203–3207.

(13) Bredt, D. S. Endogenous Nitric Oxide Synthesis: Biological Functions and Pathophysiology. *Free Radic. Res.* **1999**, *31* (6), 577–596.

(14) Friebe, A.; Koesling, D. The Function of NO-Sensitive Guanylyl Cyclase: What We Can Learn from Genetic Mouse Models. *Nitric Oxide - Biology and Chemistry*. 2009, pp 149–156.

(15) Ko, F. N.; Wu, C. C.; Kuo, S. C.; Lee, F. Y.; Teng, C. M. YC-1, a Novel Activator of Platelet Guanylate Cyclase. *Blood* **1994**, *84* (12), 4226–4233.

(16) Stasch, J.; Evgenov, O. V. *Pharmacotherapy of Pulmonary Hypertension*; Humbert, M., Ed.; 2013; Vol. 218.

(17) Cary, S. P. L.; Winger, J. A.; Marletta, M. A. Tonic and Acute Nitric Oxide Signaling through Soluble Guanylate Cyclase Is Mediated by Nonheme Nitric Oxide, ATP, and GTP. *Proc. Natl. Acad. Sci.* **2005**, *102* (37), 13064–13069.
(18) Follmann, M.; Griebenow, N.; Hahn, M. G.; Hartung, I.; Mais, F. J.; Mittendorf, J.; Sch??fer, M.; Schirok, H.; Stasch, J. P.; Stoll, F.; et al. The Chemistry and Biology of Soluble Guanylate Cyclase Stimulators and Activators. *Angew. Chemie - Int. Ed.* **2013**, *52* (36), 9442–9462.
(19) Costell, M. H.; Ancellin, N.; Bernard, R. E.; Zhao, S. F.; Upson, J. J.; Morgan, L. A.; Maniscalco, K.; Olzinski, A. R.; Ballard, V. L. T.; Herry, K.; et al. Comparison of Soluble Guanylate Cyclase Stimulators and Activators in Models of Cardiovascular Disease Associated with Oxidative Stress. *Front. Pharmacol.* **2012**, *3 JUL* (July), 1–14.
(20) Budworth, J.; Meillerais, S.; Charles, I.; Powell, K. Tissue Distribution of the Human Soluble Guanylate Cyclases. *Biochem. Biophys. Res. Commun.* **1999**, *263* (3), 696–701.
(21) Zhao, Y.; Schelvis, J.; Babcock, G.; Marletta, M. A. Identification of Histidine 105 in the Beta1 Subunit of Soluble Guanylate Cyclase as the Heme Proximal Ligand. *Biochemistry* **1998**, *37* (97), 4502–4509.
(22) Derbyshire, E. R.; Marletta, M. A. Structure and Regulation of Soluble Guanylate Cyclase. *Annu. Rev. Biochem.* **2012**, *81* (1), 533–559.
(23) Montfort, W. R.; Wales, J.; Weichsel, A. Structure and Activation of Soluble Guanylyl Cyclase, the Nitric Oxide Sensor. *Antioxidants redox Signal.* **2016**, *00* (00), 1–16.
(24) Nioche, P.; Berka, V.; Vipond, J.; Minton, N.; Tsai, A. L.; Raman, C. S. Femtomolar Sensistivity of a NO Sensor from Clostridium Botulinum. *Science (80-.).* **2004**, *306* (November), 1550–1554.
(25) Iyer, L.; Anantharaman, V.; Aravind, L. Ancient Conserved Domains Shared by Animal Soluble Guanylyl Cyclases and Bacterial Signaling Proteins. *BMC Genomics* **2003**, *4* (1), 5.
(26) Karow, D. S.; Pan, D.; Tran, R.; Pellicena, P.; Presley, A.; Mathies, R. A.; Marletta, M. A. Spectroscopic Characterization of the Soluble Guanylate Cyclase-like Heme Domains from Vibrio Cholerae and Thermoanaerobacter Tengcongensis. *Biochemistry* **2004**, *43* (31), 10203–10211.
(27) Ma, X.; Sayed, N.; Beuve, A.; van den Akker, F. NO and CO Differentially Activate Soluble Guanylyl Cyclase via a Heme Pivot-Bend Mechanism. *EMBO J.* **2007**, *26* (2), 578–588.
(28) Martin, F.; Baskaran, P.; Ma, X.; Dunten, P. W.; Schaefer, M.; Stasch, J. P.; Beuve, A.; Van Den Akker, F. Structure of Cinaciguat (BAY 58-2667) Bound to Nostoc H-NOX Domain Reveals Insights into Heme-Mimetic Activation of the Soluble Guanylyl Cyclase. *J. Biol. Chem.* **2010**, *285* (29), 22651–22657.
(29) Rao, M.; Herzik, M. A.; Iavarone, A. T.; Marletta, M. A. Nitric Oxide-Induced Conformational Changes Govern H-NOX and Histidine Kinase Interaction and Regulation in Shewanella Oneidensis. *Biochemistry* **2017**, *56* (9), 1274–1284.
(30) Zhao, Y.; Brandish, P.; Ballou, D.; Marletta, M. A. A Molecular Basis for Nitric Oxide Sensing by Soluble Guanylate Cyclase. *Proc. Natl. Acad. Sci. U. S. A.* **1999**, *96* (26), 14753–14758.
(31) Yoo, B.-K.; Lamarre, I.; Martin, J.-L.; Rappaport, F.; Negrerie, M. Motion of Proximal Histidine and Structural Allosteric Transition in Soluble Guanylate Cyclase. *Proc. Natl. Acad. Sci.* **2015**, *112* (14), E1697–E1704.

(32) Stone, J. R.; Marletta, M. A. Soluble Guanylate Cyclase from Bovine Lung: Activation with Nitric Oxide and Carbon Monoxide and Spectral Characterization of the Ferrous and Ferric States. *Biochemistry* **1994**, *33* (18), 5636–5640.
(33) Hunt, A. P.; Lehnert, N. Heme-Nitrosyls: Electronic Structure Implications for Function in Biology. *Acc. Chem. Res.* **2015**, *48* (7), 2117–2125.
(34) Herzik, M. A.; Jonnalagadda, R.; Kuriyan, J.; Marletta, M. A. Structural Insights into the Role of Iron–Histidine Bond Cleavage in Nitric Oxide-Induced Activation of H-NOX Gas Sensor Proteins. *Proc. Natl. Acad. Sci.* **2014**, *111* (40), E4156–E4164.
(35) Baskaran, P.; Heckler, E. J.; van den Akker, F.; Beuve, A. Identification of Residues in the Heme Domain of Soluble Guanylyl Cyclase That Are Important for Basal and Stimulated Catalytic Activity. *PLoS One* **2011**, *6* (11), e26976.
(36) Baskaran, P.; Heckler, E. J.; Van Den Akker, F.; Beuve, A. Aspartate 102 in the Heme Domain of Soluble Guanylyl Cyclase Has a Key Role in No Activation. *Biochemistry* **2011**, *50* (20), 4291–4297.
(37) Henry, J. T.; Crosson, S. Ligand-Binding PAS Domains in a Genomic, Cellular, and Structural Context. *Annu. Rev. Microbiol.* **2011**, *65* (June), 261–286.
(38) Möglich, A.; Ayers, R. A.; Moffat, K. Structure and Signaling Mechanism of Per-ARNT-Sim Domains. *Structure*. 2009, pp 1282–1294.
(39) Ma, X.; Sayed, N.; Baskaran, P.; Beuve, A.; Van Den Akker, F. PAS-Mediated Dimerization of Soluble Guanylyl Cyclase Revealed by Signal Transduction Histidine Kinase Domain Crystal Structure. *J. Biol* **2007**, *283* (2), 1167–1178.
(40) Zhao, Y.; Marletta, M. A. Localization of the Heme Binding Region in Soluble Guanylate Cyclase. *Biochemistry* **1997**, *36* (50), 15959–15964.
(41) Fritz, B.; Roberts, S.; Ahmed, A.; Breci, L.; Li, W.; Weichsel, A.; Brailey, J.; Wysocki, V.; Tama, F.; Montfort, W. Molecular Model of a Soluble Guanylyl Cyclase Fragment Determined by Small-Angle X-Ray Scattering and Chemical Cross-Linking. *Biochemistry* **2013**, *52* (9), 1568–1582.
(42) Sarkar, A.; Dai, Y.; Haque, M. M.; Seeger, F.; Ghosh, A.; Garcin, E. D.; Montfort, W. R.; Hazen, S. L.; Misra, S.; Stuehr, D. J. Heat Shock Protein 90 Associates with the Per-Arnt-Sim Domain of Heme-Free Soluble Guanylate Cyclase. *J. Biol. Chem.* **2015**, *290* (35), 21615–21628.
(43) Purohit, R.; Weichsel, A.; Montfort, W. R. Crystal Structure of the Alpha Subunit PAS Domain from Soluble Guanylyl Cyclase. *Protein Sci.* **2013**, *22* (10), 1439–1444.
(44) Ma, X.; Beuve, A.; van den Akker, F. Crystal Structure of the Signaling Helix Coiled-Coil Domain of the Beta1 Subunit of the Soluble Guanylyl Cyclase. *BMC Struct. Biol.* **2010**, *10* (2).
(45) Campbell, M. G.; Underbakke, E. S.; Potter, C. S.; Carragher, B.; Marletta, M. A. Single-Particle EM Reveals the Higher-Order Domain Architecture of Soluble Guanylate Cyclase. *Proc. Natl. Acad. Sci.* **2014**, *111* (8), 2960–2965.
(46) Underbakke, E. S.; Iavarone, A. T.; Chalmers, M. J.; Pascal, B. D.; Novick, S.; Griffin, P. R.; Marletta, M. A. Nitric Oxide-Induced Conformational Changes in Soluble Guanylate Cyclase. *Cell Struct.* **2014**, *22* (4), 602–611.
(47) Linder, J. U. Class III Adenylyl Cyclases: Molecular Mechanisms of Catalysis and Regulation. *Cellular and Molecular Life Sciences*. 2006, pp 1736–1751.
(48) Winger, J. A.; Derbyshire, E. R.; Lamers, M. H.; Marletta, M. A.; Kuriyan, J. The Crystal Structure of the Catalytic Domain of a Eukaryotic Guanylate Cyclase. *BMC Struct. Biol.*

2008, *8*, 42–52.
(49) Allerston, C. K.; von Delft, F.; Gileadi, O. Crystal Structures of the Catalytic Domain of Human Soluble Guanylate Cyclase. *PLoS One* **2013**, *8* (3), e57644.
(50) Seeger, F.; Quintyn, R.; Tanimoto, A.; Williams, G. J.; Tainer, J. A.; Wysocki, V. H.; Garcin, E. D. Interfacial Residues Promote an Optimal Alignment of the Catalytic Center in Human Soluble Guanylate Cyclase: Heterodimerization Is Required but Not Sufficient for Activity. *Biochemistry* **2014**, *53* (13), 2153–2165.
(51) Derbyshire, E. R.; Fernhoff, N. B.; Deng, S.; Marletta, M. A. Nucleotide Regulation of Soluble Guanylate Cyclase Substrate Specificity. *Biochemistry* **2009**, *48* (31), 7519–7524.
(52) Sürmeli, N. B.; Müskens, F. M.; Marletta, M. A. The Influence of Nitric Oxide on Soluble Guanylate Cyclase Regulation by Nucleotides. *J. Biol. Chem.* **2015**, *290* (25), 15570–15580.
(53) Childers, K. C.; Garcin, E. D. Nitric Oxide Structure / Function of the Soluble Guanylyl Cyclase Catalytic Domain. *Nitric Oxide* **2018**, *77*, 53–64.
(54) Kang, Y.; Liu, R.; Wu, J.-X.; Chen, L. Structural Insights into the Mechanism of Human Soluble Guanylate Cyclase. *Nature* **2019**.
(55) Vercellino, I.; Rezabkova, L.; Olieric, V.; Polyhach, Y.; Weinert, T.; Kammerer, R. A.; Jeschke, G.; Korkhov, V. M. Role of the Nucleotidyl Cyclase Helical Domain in Catalytically Active Dimer Formation. *Proc. Natl. Acad. Sci.* **2017**, *114* (46), E9821–E9828.
(56) Winger, J. A.; Marletta, M. A. Expression and Characterization of the Catalytic Domains of Soluble Guanylate Cyclase: Interaction with the Heme Domain. *Biochemistry* **2005**, *44* (10), 4083–4090.
(57) Stone, J. R.; Marletta, M. A. Spectral and Kinetic Studies on the Activation of Soluble Guanylate Cyclase by Nitric Oxide. *Biochemistry* **1996**, *35* (4), 1093–1099.
(58) Martin, E.; Berka, V.; Tsai, A.-L.; Murad, F. Soluble Guanylyl Cyclase: The Nitric Oxide Receptor. *Methods Enzymol.* **2005**, *396* (1996), 478–492.
(59) Martin, E.; Berka, V.; Sharina, I.; Tsai, A. L. Mechanism of Binding of NO to Soluble Guanylyl Cyclase: Implication for the Second NO Binding to the Heme Proximal Site. *Biochemistry* **2012**, *51* (13), 2737–2746.
(60) Makino, R.; Matsuda, H.; Obayashiţ, E.; Shiro, Y.; Iizuka, T.; Hori, H. EPR Characterization of Axial Bond in Metal Center of Native and Cobalt-Substituted Guanylate Cyclase. *J. Biol. Chem.* **1999**, *274* (12), 7714–7723.
(61) Martin, E.; Berka, V.; Bogatenkova, E.; Murad, F.; Tsai, A. L. Ligand Selectivity of Soluble Guanylyl Cyclase: Effect of the Hydrogen-Bonding Tyrosine in the Distal Heme Pocket on Binding of Oxygen, Nitric Oxide, and Carbon Monoxide. *J. Biol. Chem.* **2006**, *281* (38), 27836–27845.
(62) Dierks, E. A.; Hu, S.; Vogel, K. M.; Yu, A. E.; Spiro, T. G.; Burstyn, J. N. Demonstration of the Role of Scission of the Proximal Histidine-Iron Bond in the Activation of Soluble Guanylyl Cyclase through Metalloporphyrin Substitution Studies. *J. Am. Chem. Soc.* **1997**, *119*, 7316–7323.
(63) Bellamy, T. C.; Wood, J.; Garthwaite, J. On the Activation of Soluble Guanylyl Cyclase by Nitric Oxide. *Proc. Natl. Acad. Sci.* **2002**, *99* (1), 507–510.
(64) Ballou, D. P.; Zhao, Y.; Brandish, P. E.; Marletta, M. A. Revisiting the Kinetics of Nitric Oxide (NO) Binding to Soluble Guanylate Cyclase: The Simple NO-Binding Model Is Incorrect. *Proc. Natl. Acad. Sci. U. S. A.* **2002**, *99* (19), 12097–12101.

(65) Lawson, D. M.; Stevenson, C. E.; Andrew, C. R.; Eady, R. R. Unprecedented Proximal Binding of Nitric Oxide to Heme: Implications for Guanylate Cyclase. *EMBO J.* **2000**, *19* (21), 5661–5671.
(66) Hespen, C. W.; Bruegger, J. J.; Phillips-Piro, C. M.; Marletta, M. A. Structural and Functional Evidence Indicates Selective Oxygen Signaling in Caldanaerobacter Subterraneus H-NOX. *ACS Chem. Biol.* **2016**, *11* (8), 2337–2346.
(67) Guo, Y.; Suess, D. L. M.; Herzik, M. A.; Iavarone, A. T.; Britt, R. D.; Marletta, M. A. Regulation of Nitric Oxide Signaling by Formation of a Distal Receptor–Ligand Complex. *Nat. Chem. Biol.* **2017**, *13* (12), 1216–1221.
(68) Négrerie, M.; Bouzhir, L.; Martin, J. L.; Liebl, U. Control of Nitric Oxide Dynamics by Guanylate Cyclase in Its Activated State. *J. Biol. Chem.* **2001**, *276* (50), 46815–46821.
(69) Yoo, B. K.; Lamarre, I.; Martin, J. L.; Negrerie, M. Quaternary Structure Controls Ligand Dynamics in Soluble Guanylate Cyclase. *J. Biol. Chem.* **2012**, *287* (9), 6851–6859.
(70) Bellamy, T. C.; Griffiths, C.; Garthwaite, J. Differential Sensitivity of Guanylyl Cyclase and Mitochondrial Respiration to Nitric Oxide Measured Using Clamped Concentrations. *J. Biol. Chem.* **2002**, *277* (35), 31801–31807.
(71) Russwurm, M.; Koesling, D. NO Activation of Guanylyl Cyclase. *EMBO J.* **2004**, *23* (22), 4443–4450.
(72) Derbyshire, E. R.; Marletta, M. A. Butyl Isocyanide as a Probe of the Activation Mechanism of Soluble Guanylate Cyclase: Investigating the Role of Non-Heme Nitric Oxide. *J. Biol. Chem.* **2007**, *282* (49), 35741–35748.
(73) Fernhoff, N. B.; Derbyshire, E. R.; Marletta, M. A. A Nitric Oxide/Cysteine Interaction Mediates the Activation of Soluble Guanylate Cyclase. *Proc. Natl. Acad. Sci.* **2009**, *106* (51), 21602–21607.
(74) Stasch, J. P.; Becker, E. M.; Alonso-Alija, C.; Apeler, H.; Dembowsky, K.; Feurer, A.; Gerzer, R.; Minuth, T.; Perzborn, E.; Pleiss, U.; et al. NO-Independent Regulatory Site on Soluble Guanylate Cyclase. *Nature* **2001**, *410* (6825), 212–215.
(75) Lamothe, M.; Chang, F. J.; Balashova, N.; Shirokov, R.; Beuve, A. Functional Characterization of Nitric Oxide and YC-1 Activation of Soluble Guanylyl Cyclase: Structural Implication for the YC-1 Binding Site? *Biochemistry* **2004**, *43* (11), 3039–3048.
(76) Martin, E.; Czarnecki, K.; Jayaraman, V.; Murad, F.; Kincaid, J. Resonance Raman and Infrared Spectroscopic Studies of High-Output Forms of Human Soluble Guanylyl Cyclase. *J. Am. Chem. Soc.* **2005**, *127*, 4625–4631.
(77) Yazawa, S.; Tsuchiya, H.; Hori, H.; Makino, R. Functional Characterization of Two Nucleotide-Binding Sites in Soluble Guanylate Cyclase. *J. Biol. Chem.* **2006**, *281* (31), 21763–21770.
(78) Hu, X.; Murata, L. B.; Weichsel, A.; Brailey, J. L.; Roberts, S. A.; Nighorn, A.; Montfort, W. R. Allostery in Recombinant Soluble Guanylyl Cyclase from Manduca Sexta. *J. Biol. Chem.* **2008**, *283* (30), 20968–20977.
(79) Ibrahim, M.; Derbyshire, E. R.; Marletta, M. A.; Spiro, T. G. Probing Soluble Guanylate Cyclase Activation by CO and YC-1 Using Resonance Raman Spectroscopy. *pubs.acs.org/Biochemistry Biochem.* **2010**, *49*, 3815–3823.
(80) Ibrahim, M.; Derbyshire, E. R.; Soldatova, A. V; Marletta, M. A.; Spiro, T. G. Soluble Guanylate Cyclase Is Activated Differently by Excess NO and by YC-1: Resonance Raman Spectroscopic Evidence. *Biochemistry* **2010**, *49*, 4864–4871.
(81) Yoo, B.-K.; Lamarre, I.; Martin, J.-L.; Negrerie, M. Quaternary Structure Controls Ligand

Dynamics in Soluble Guanylate Cyclase. *J. Biol. Chem.* **2012**, *287* (9), 6851–6859.
(82) Purohit, R.; Fritz, B. G.; The, J.; Issaian, A.; Weichsel, A.; David, C. L.; Campbell, E.; Hausrath, A. C.; Rassouli-Taylor, L.; Garcin, E. D.; et al. YC-1 Binding to the β Subunit of Soluble Guanylyl Cyclase Overcomes Allosteric Inhibition by the α Subunit. *Biochemistry* **2014**, *53* (1), 101–114.
(83) Wales, J. A.; Chen, C.; Breci, L.; Weichsel, A.; Bernier, S. G.; Sheppeck II, J. E.; Solinga, R.; Nakai, T.; Renhowe, P. A.; Jung, J.; et al. Discovery of Stimulator Binding to a Conserved Pocket in the Heme Domain of Soluble Guanylyl Cyclase. *J. Biol. Chem.* **2018**, *293* (5), 1850–1864.
(84) Bellamy, T. C.; Garthwaite, J. Sub-Second Kinetics of the Nitric Oxide Receptor, Soluble Guanylyl Cyclase, in Intact Cerebellar Cells. *J. Biol. Chem.* **2001**, *276* (6), 4287–4292.
(85) Hall, C. N.; Garthwaite, J. What Is the Real Physiological NO Concentration in Vivo? *Nitric Oxide - Biol. Chem.* **2009**, *21* (2), 92–103.
(86) Roy, B.; Halvey, E. J.; Garthwaite, J. An Enzyme-Linked Receptor Mechanism for Nitric Oxide-Activated Guanylyl Cyclase. *J. Biol. Chem.* **2008**, *283* (27), 18841–18851.
(87) Roy, B.; Garthwaite, J. Nitric Oxide Activation of Guanylyl Cyclase in Cells Revisited. *Proc. Natl. Acad. Sci. U. S. A.* **2006**, *103* (32), 12185–12190.
(88) Batchelor, A. M.; Bartus, K.; Reynell, C.; Constantinou, S.; Halvey, E. J.; Held, K. F.; Dostmann, W. R.; Vernon, J.; Garthwaite, J. Exquisite Sensitivity to Subsecond, Picomolar Nitric Oxide Transients Conferred on Cells by Guanylyl Cyclase-Coupled Receptors. *Proc. Natl. Acad. Sci.* **2010**, *107* (51), 22060–22065.
(89) Wood, K. C.; Batchelor, A. M.; Bartus, K.; Harris, K. L.; Garthwaite, G.; Vernon, J.; Garthwaite, J. Picomolar Nitric Oxide Signals from Central Neurons Recorded Using Ultrasensitive Detector Cells. *J. Biol. Chem.* **2011**, *286* (50), 43172–43181.
(90) Mo, E.; Amin, H.; Bianco, I. H.; Garthwaite, J. Kinetics of a Cellular Nitric Oxide/CGMP/Phosphodiesterase-5 Pathway. *J. Biol. Chem.* **2004**, *279* (25), 26149–26158.
(91) Halvey, E. J.; Vernon, J.; Roy, B.; Garthwaite, J. Mechanisms of Activity-Dependent Plasticity in Cellular Nitric Oxide-CGMP Signaling. *J. Biol. Chem.* **2009**, *284* (38), 25630–25641.
(92) Gray, J. M.; Karow, D. S.; Lu, H.; Chang, A. J.; Chang, J. S.; Ellis, R. E.; Marletta, M. A.; Bargmann, C. I. Oxygen Sensation and Social Feeding Mediated by a *C. Elegans* Guanylate Cyclase Homologue. *Nature* **2004**, *430* (6997), 317–322.
(93) Morton, D. B. Atypical Soluble Guanylyl Cyclases in Drosophila Can Function as Molecular Oxygen Sensors. *J. Biol. Chem.* **2004**, *279* (49), 50651–50653.
(94) Margulis, A.; Sitaramayya, A. Rate of Deactivation of Nitric Oxide-Stimulated Soluble Guanylate Cyclase: Influence of Nitric Oxide Scavengers and Calcium. *Biochemistry* **2000**, *39*, 1034–1039.
(95) Mergia, E.; Friebe, A.; Dangel, O.; Russwurm, M.; Koesling, D. Spare Guanylyl Cyclase NO Receptors Ensure High NO Sensitivity in the Vascular System. *J. Clin. Invest.* **2006**, *116* (6), 1731–1737.
(96) Bojkowska, K.; Santoni De Sio, F.; Barde, I.; Offner, S.; Verp, S.; Heinis, C.; Johnsson, K.; Trono, D. Measuring in Vivo Protein Half-Life. *Chem. Biol.* **2011**, *18* (6), 805–815.
(97) Ritchie, R. H.; Drummond, G. R.; Sobey, C. G.; De Silva, T. M.; Kemp-Harper, B. K. The Opposing Roles of NO and Oxidative Stress in Cardiovascular Disease. *Pharmacol. Res.* **2016**, *116*, 57–69.

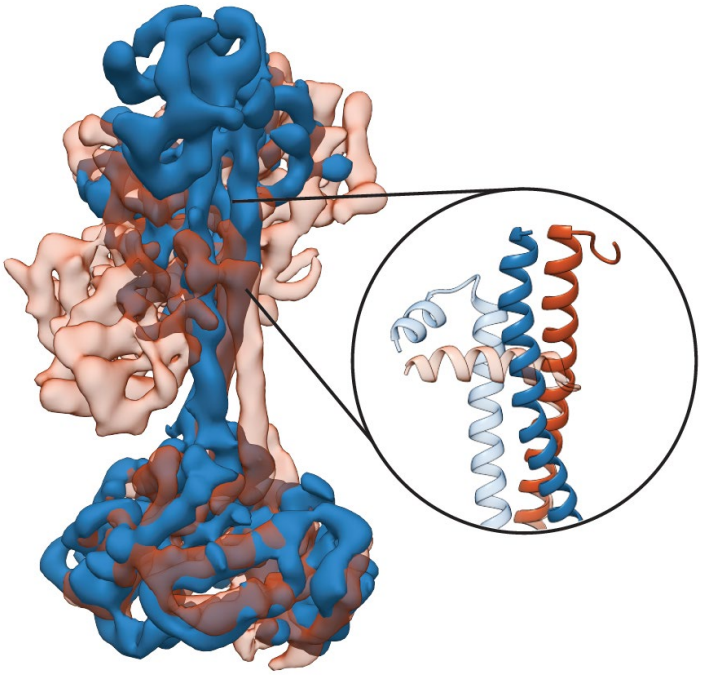
(98) Ghosh, A.; Koziol-White, C. J.; Asosingh, K.; Cheng, G.; Ruple, L.; Groneberg, D.; Friebe, A.; Comhair, S. A. A.; Stasch, J.-P.; Panettieri, R. A.; et al. Soluble Guanylate Cyclase as an Alternative Target for Bronchodilator Therapy in Asthma. *Proc. Natl. Acad. Sci.* **2016**, *113* (17), E2355–E2362.
(99) Beuve, A. Thiol-Based Redox Modulation of Soluble Guanylyl Cyclase, the Nitric Oxide Receptor. *Antioxid. Redox Signal.* **2017**, *26* (3), 137–149.
(100) Stasch, J. P.; Pacher, P.; Evgenov, O. V. Soluble Guanylate Cyclase as an Emerging Therapeutic Target in Cardiopulmonary Disease. *Circulation* **2011**, *123* (20), 2263–2273.
(101) Sayed, N.; Baskaran, P.; Ma, X.; van den Akker, F.; Beuve, A. Desensitization of Soluble Guanylyl Cyclase, the NO Receptor, by S-Nitrosylation. *Proc. Natl. Acad. Sci. U. S. A.* **2007**, *104* (30), 12312–12317.
(102) Beuve, A.; Wu, C.; Cui, C.; Liu, T.; Jain, M. R.; Huang, C.; Yan, L.; Kholodovych, V.; Li, H. Identification of Novel S-Nitrosation Sites in Soluble Guanylyl Cyclase, the Nitric Oxide Receptor. *J. Proteomics* **2016**, *138*, 40–47.
(103) Surmeli, N. B.; Marletta, M. A. Insight into the Rescue of Oxidized Soluble Guanylate Cyclase by the Activator Cinaciguat. *ChemBioChem* **2012**, *13* (7), 977–981.
(104) Hobbs, A. J. Soluble Guanylate Cyclase, Emerging Therapeutic Targets. *Emerg. Ther. Targets* **2000**, *4* (6), 735–749.
(105) Alexandre, E. C.; Leiria, L. O.; Silva, F. H.; Mendes-Silvério, C. B.; Calmasini, F. B.; Paula, A.; Davel, C.; Mónica, F. Z.; De Nucci, G.; Antunes, E. Soluble Guanylyl Cyclase (SGC) Degradation and Impairment of Nitric Oxide-Mediated Responses in Urethra from Obese Mice: Reversal by the SGC Activator BAY 60-2770. *J. Pharmacol. Exp. Ther. J Pharmacol Exp Ther* **2014**, *349*, 2–9.
(106) Fernhoff, N. B.; Derbyshire, E. R.; Underbakke, E. S.; Marletta, M. A. Heme-Assisted S-Nitrosation Desensitizes Ferric Soluble Guanylate Cyclase to Nitric Oxide. *J. Biol. Chem.* **2012**, *287* (51), 43053–43062.

CHAPTER 2

ALLOSTERIC ACTIVATION OF THE NITRIC OXIDE RECEPTOR SOLUBLE GUANYLATE CYCLASE MAPPED BY CRYO-ELECTRON MICROSCOPY

Summary

Soluble guanylate cyclase (sGC) is the primary receptor for nitric oxide (NO) in mammalian nitric oxide signaling. We determined structures of full-length *Manduca sexta* sGC in both inactive and active states using cryo-electron microscopy. NO and the sGC-specific stimulator YC-1 induce a 71° rotation of the heme-binding β H-NOX and PAS domains. Repositioning of the β H-NOX domain leads to a straightening of the coiled-coil domains, which, in turn, use the motion to move the catalytic domains into an active conformation. YC-1 binds directly between the β H-NOX domain and the two CC domains. The structural elongation of the particle observed in cryo-EM was corroborated in solution using small angle X-ray scattering (SAXS). These structures delineate the endpoints of the allosteric transition responsible for the major cyclic GMP-dependent physiological effects of NO.



Portions of this chapter have been reported in a previous publication.[1]

Introduction

Nitric oxide (NO) is a critical primary signaling molecule in eukaryotic organisms.[2,3] Cells detect this diatomic gas through the NO receptor soluble guanylate cyclase (sGC), which is activated by NO to catalyze the cyclization of 5′-guanosine triphosphate (GTP) to 3′,5′-cyclic guanosine monophosphate (cGMP).[4–9] The NO signaling pathway, with sGC as a central component, plays crucial roles in the cardiovascular and neurological systems in mammals.[10–12] Furthermore, aberrations in these signaling pathways can lead to pathologies that include various forms of hypertension, cardiovascular disease, and neurodegeneration.[13–17]

sGC is a heterodimer composed of two subunits, denoted α and β, that are each composed of four domains: an N-terminal heme nitric oxide/oxygen (H-NOX) domain, a Per/Arnt/Sim (PAS)-like domain, a coiled-coil (CC) domain, and a catalytic (CAT) domain.[7,8] Although each subunit contains an H-NOX domain, only the β H-NOX domain binds a heme cofactor, with direct ligation occurring through a conserved histidine residue.[18] The CC domains, together with the PAS domains, are thought to form a structured assembly upon dimerization of sGC.[19] The CAT domains form a wreath-like structure with the active site at the dimer interface.[7,20]

Biochemical aspects of sGC activation by NO have been studied in great detail. Without a ligand bound to the β H-NOX domain, sGC has a low basal activity. A stoichiometric equivalent of NO relative to the sGC heterodimer results in the cleavage of the proximal histidine-iron bond and the formation of a distal five-coordinate ferrous nitrosyl enzyme with 15% of maximal activity.[6,21–23] This low-activity state of sGC will be referred to here as the 1-NO state; importantly, the K_D of the ferrous nitrosyl heme is 1.2×10^{-12} M thus in this state NO remains bound to the heme.[24] The activity of the 1-NO state can be increased to a maximally-active state either by the addition of excess NO (xsNO), or by addition of small-molecule stimulators.[25] The benzylindazole compound YC-1 was the first reported small-molecule sGC stimulator.[26] It was identified in a screen of compounds that inhibit platelet aggregation, one of the physiological responses of sGC activation. The molecular mechanisms by which NO and small molecule stimulator binding leads to enzyme activation remain unclear, despite the fact that the sGC-targeted drug Adempas® discovered through a small molecule screen based on the YC-1 scaffold was approved by the FDA in 2013.

Visualization of the molecular steps in sGC activation has been a longstanding challenge. Crystal structures of individual truncated domains from various homologues of sGC have been reported.[27–30] Additionally, structures of related heterodimeric and multidomain proteins have provided insight into the higher order connectivity of sGC. The heterodimeric catalytic domain from *Homo sapiens* was solved in an inactive conformation.[31,32] Structures of membrane-bound adenylate cyclases have been solved that contain catalytic domains as well as portions of or the entire CC domains, helping to orient the C-terminal domains of sGC.[33,34] However, the precise quaternary structural arrangement of domains in the N-terminal portion of sGC is not known. Consequently, the mechanism by which NO and small-molecule stimulators couple binding through the protein for activation are poorly understood.

In the absence of a high-resolution full-length structure, previous work has used alternative methods in attempts to understand interdomain interactions and how small molecules activate sGC.

Crosslinking experiments and negative stain electron microscopy support the hypothesis that sGC is a flexible dumbbell-shaped particle, in which the CC domains serve to connect the H-NOX and PAS domains on one end of the dumbbell to the CAT domain on the other.[19,35] Hydrogen deuterium exchange mass spectrometry (HDX-MS) has implicated the linker region between the PAS domains and the CC domains as critical in the activation mechanism, as both regions change in H/D exchange upon NO binding to sGC.[36] Truncations of sGC have suggested that the β H-NOX domain directly inhibits the CAT domains.[37] Point mutagenesis was used to identify several residues thought to transmit the ligand occupancy of the gas-binding H-NOX domain to the catalytic domains, including β D102 and β D106.[38,39] None of these methods, however, afford sufficient resolution to discern domain organization and interdomain communication for the full-length protein.

Here, we report the first full-length structures of sGC with and without activating ligands using cryoelectron microscopy (cryo-EM) to 5.1 and 5.8 Å resolution, respectively. We find that sGC adopts two dramatically different conformations, revealing a large-scale global conformational change in response to NO and YC-1. Density consistent with YC-1 is observed in the active state map directly between the β H-NOX and both CC domains. We also show the functional relationship of NO binding to sGC on overall organization of sGC in solution by small angle X-ray scattering (SAXS). By visualizing the quaternary structural changes that activate this NO receptor, we can now answer the long-standing question as to how the ligand binding in the regulatory domains of sGC is communicated to the catalytic domains.

Materials and Methods

Key Resource Table

Key Resources Table				
Reagent type (species) or resource	Designation	Source or reference	Identifiers	Additional information
gene (*Manduca sexta*)	*Ms* sGC α1	Integrated DNA Technologies	Genbank: AF062750	Uniprot: O77105
gene (*Manduca sexta*)	*Ms* sGC β1	Integrated DNA Technologies	Genbank: AF062751	Uniprot: O77106
strain, strain background (*E. coli*)	XL1-Blue	UC Berkeley MacroLab		Cloning strain
strain, strain background (*E. coli*)	DH10-Bac	UC Berkeley MacroLab		Transposition Strain
recombinant DNA reagent	pFastBac (plasmid)	Thermo Fisher		Donor Plasmid

recombinant DNA reagent	pFastBac_Ms_α1_His6	This paper		
recombinant DNA reagent	pFastBac_Ms_β1	This paper		
Sequence-based regent	Primers			See primer table in materials and methods
cell line (*Spodoptera frugiperda*)	SF9 cells	UC Berkeley Cell Culture Facility		Expression Strain
Commercial Assay or Kit	Cellfectin II	Thermo Fisher		Transfection reagent
Transfected Construct (*Manduca sexta)*	bMON14272_ Ms_α1_His6 (bacmid)	This paper		Transfected construct
Transfected Construct (*Manduca sexta)*	bMON14272_ Ms_β1_His6	This paper		Transfected construct
Chemical compound	DEA NONOate	Cayman Chemical	Item No: 82100	CAS: 372965-00-9
Chemical compound	PROLI NONOate	Cayman Chemical	Item No: 82145	CAS: 178948-42-0
Commercial Assay or Kit	cGMP ELISA	Enzo Life Science	ADI-901-013	

Materials

All chemicals were purchased from commercial vendors and used without further purification, unless otherwise noted. Primers and gBlocks were obtained from Integrated DNA Technologies (Coralville, IA). Gibson Master Mix was purchased from New England Biolabs (Ipswich, MA). PrimeSTAR Max DNA polymerase and TALON superflow resin was purchased from Takara Bio USA (Mountain View, CA). DNA purification kits were purchased from Qiagen (Germantown, MD). ExCell-420 media, sodium dithionite ($Na_2S_2O_4$), DNAse I from bovine pancreas, β-mercaptoethanol, sodium phosphate, dibasic (Na_2HPO_4), sodium carbonate (Na_2CO_3), potassium phosphate (K_2HPO_4), guanosine 5′-triphosphate sodium salt (GTP) (>95%, HPLC), and zinc acetate ($Zn(CH_3CO_2)_2$) were purchased from Millipore Sigma (Burlington, MA). 4-(2-aminoethyl)benzenesulfonyl fluoride hydrochloride (AEBSF) was purchased from Research

Products International (Mount Prospect, IL). Isopropyl β-D-1-thiogalactopyranoside (IPTG), 4-(2-hydroxyethyl)piperazine-1-ethanesulfonic acid (HEPES), and the Pierce BCA Protein Assay Kit were purchased from Thermo Fisher Scientific (Waltham, MA). Benzamidine, sodium chloride (NaCl), magnesium chloride ($MgCl_2$), and glycerol were purchased from VWR Life Science (Visalia, CA). Dithiothreitol (DTT) was purchased from Bachem (Bubendorf, Switzerland). Imidazole was purchased from Oakwood Chemical (West Columbia, SC). Vivaspin spin concentrators were purchased from Sartorius (Concord, CA). BioSpin 6 desalting columns were purchased from BioRad (Hercules, CA). Diethylamine NONOate (DEA NONOate), PROLI NONOate, diethylamintriamine NONOate (DETA NONOate), and YC-1 were purchased from Cayman Chemical Company (Ann Arbor, MI). Carbon monoxide (CO, >99%) gas was purchased from Praxair Inc. (Danbury, CT).

Construction of Plasmids

Manduca sexta sGC α1 (Uniprot: O77105) and β1(Uniprot: O77106) genes were purchased as gBlocks from IDT with a C-terminal 6x His tag on the α1 gene with Golden Gate cloning sites. The gBlocks were subcloned into a Golden Gate entry vector (gift of the Tullman-Ercek lab, Northwestern University). PCR was used to add regions of homology to pFastBac (Bac-to-Bac Baculovirus Expresion System, Thermo Fisher Scientific), and Gibson Assembly was performed to generate pFastBac_*Ms*_sGC_α1_His6 and pFastBac_*Ms*_sGC_β1. The genes were transposed into a baculovirus bacmid using DH10Bac-GFP cells. The bacmid was isolated, validated, and then transfected into SF9 cells (Berkeley Cell Culture Facility) to generate recombinant baculovirus for both genes.

Name	Sequence (5' → 3')	Target	Purpose
oBGH 130	GGAGATAATTAAAATGATAACCATCTCGC	pFastBac	Sequencing
oBGH 131	TTTATTTGTGAAATTTGTGATGCTATTGC	pFastBac	Sequencing
oBGH 297	CGATGGAGTACGAAGCGAATTTCGTATG	Ms_α	Sequencing
oBGH 168	GCCTGCACGATCATCTCGGGAC	Ms_β	Sequencing
oBGH 195	CTGCCAAGCTCCAGGAACACC ATGACGTGTCCATTCCGTCGTG	Ms_α (F)	Cloning
oBGH 196	CCTCGAGACTGCAGGCTCTAGA TTAGTGATGGTGATGGTGATG	Ms_α (R)	Cloning
oBGH 197	GCTCCGGGAGCGGACACC ATGTACGGGTTTGTGAACTATGCCC	Ms_β (F)	Cloning
oBGH 198	CCCCACAAAGGCTGCCAT TTAATGGATCTTCCTGGTGAGGAACCAG	Ms_β (R)	Cloning
oBGH 199	CATCACCATCACCATCACTAATCTAGAGCCTGCAGTC TCGAGG	pFastBac (F)	Cloning
oBGH 200	CACGACGGAATGGACACGTCATGGTGTTCCTGGAGC TTGGCAG	pFastBac (R)	Cloning
oBGH 201	CTGGTTCCTCACCAGGAAGATCCATTAAATGGCAGC CTTTGTGGGG	pFastBac (F)	Cloning

oBGH 202	GGGCATAGTTCACAAACCCGTACATGGTGTCCGCTCC CGGAGC	pFastB ac (R)	Cloning

Protein Expression and Purification
SF9 cells were maintained in monolayer and in suspension in ExCell-420 media at 27 °C (cells were shaken at 135 rpm). The recombinant baculovirus was amplified until the titer was greater than 1×10^8 cfu/mL. Five liters of SF9 cells were coinfected with 50 mL of amplified virus per liter and allowed to express for 72 hours. Cells were spun down at 4300 *g* for 20 minutes, snap frozen in liquid nitrogen, and stored at –80 °C.

The following steps were performed at 4 °C. Cell pellets were thawed in ice water and resuspended in lysis buffer: Buffer A (50 mM Na_2HPO_4, pH 8.0, 200 mM NaCl, 1 mM imidazole, 1 mM benzamidine, 5% (v/v) glycerol, 0.22 μm filtered) supplemented with 10 mM benzamidine, 1 mM AEBSF, 0.5 mg/mL bovine DNAse I, and 5 mM β-mercaptoethanol. Cells were lysed in a bead beater (BioSpec) with 0.5 mm glass beads. The resulting cell lysate was clarified by spinning at 4,300 *g* for 5 minutes, followed by spinning at 158,000 *g* for 2 hours (Ti45 rotor, Beckman Coulter). The lysate was passed through a column containing 2 mL TALON superflow at 0.5 mL/min, and the flow-through was collected. The column was washed with 15 column volumes of Buffer A supplemented with 5 mM β-mercaptoethanol at 0.5 mL/min. The protein was eluted with 10 CV of Buffer B (50 mM Na_2HPO_4, pH 8.0, 200 mM NaCl, 150 mM imidazole, 1 mM benzamidine, 5% (v/v) glycerol, 0.22 μm filtered) supplemented with 5 mM β- mercaptoethanol. Fractions with yellow color were concentrated to < 2 mL using a 30,000 molecular weight cutoff spin concentrator, supplemented with 5 mM DTT and 5 mM EDTA, and stored overnight. Next, the sample was diluted to 9 mL with Buffer C (25 mM triethanolamine, 25 mM NaCl, 5 mM DTT, 5% (v/v) glycerol, 0.22 μm filtered), and applied to a POROS HQ2 anion exchange column (Thermo Fisher Scientific). The column was washed with 3 CV of Buffer C, and then a gradient to 50% Buffer D (25 mM triethanolamine, 750 mM NaCl, 5 mM DTT, 5% (v/v) glycerol, 0.22 μm filtered) was established over 17 CV at 0.5 mL/min. Fractions with purified sGC were concentrated to 5–50 μM and stored in liquid nitrogen. A typical yield for this expression and purification procedure is 100 μg sGC per liter of insect cells.

Intact Protein Mass Spectrometry
Purified proteins were buffer-exchanged into 25 mM HEPES, pH 7.5, 25 mM NaCl using three rounds of dilution and concentrations in a Vivaspin 500 (30,000 MWCO) spin concentrator. Samples were filtered with a 0.22 μM spin filter (Millipore). The final sample concentration was approximately 5 μM. Samples were separated with an Agilent 1200 series high pressure liquid chromatography (HPLC) system over a ProSwift column (ThermoFisher Scientific), and subsequently analyzed by an Agilent 6224 time-of-flight (TOF) mass spectrometer with a Turbospray ion source in positive ion mode.

Absorption Spectroscopy
Samples were reduced in an anaerobic glove bag (Coy) with 5 mM $Na_2S_2O_4$ for 15 minutes, and then desalted with a BioSpin6 column equilibrated with Buffer E (50 mM HEPES, pH 7.5, 150 mM NaCl, 5% (v/v) glycerol, 0.22 μm filtered). CO-saturated buffer (950 μM CO) was prepared by sparging 3 mL of anaerobic Buffer E for 15 minutes in a Reacti-Vial (Thermo Fisher Scientific).

CO was added to the sample to achieve a final concentration of 425 μM. Nitric oxide was added to a concentration of 500 μM by addition of DEA NONOate, based on 1.5 moles of NO released per mole of NONOate. Protein-ligand complexes were incubated for 15 minutes at room temperature to establish equilibrium, and no further spectral changes were observed after this time. Samples were placed in a septum-sealed 1 cm pathlength quartz cuvette inside the glove bag, and UV–Vis spectra were recorded on a Cary 300 spectrophotometer (Agilent Technologies).

Activity Assays and Quantification

Steady-state kinetics for *Ms* sGC were measured by quantifying the amount of cGMP produced in duplicate endpoint activity assays, performed in at least biological triplicate. Samples were reduced in an anaerobic glove bag with 5 mM $Na_2S_2O_4$ for 15 minutes, and then desalted with a BioSpin6 column equilibrated with Buffer E. *Ms* sGC with 1-NO and xsNO were prepared by first adding PROLI NONOate to 50 μM, based on 2 moles of NO released per mole of PROLI NONOate. This sample was then buffer exchanged into Buffer E through a BioSpin6 column to generate the 1-NO state. To generate the xsNO state, PROLI NONOate was added back to a portion of the 1-NO sample and allowed to equilibrate for 5 minutes. YC-1 was added from a 100x stock solution in DMSO to a final concentration of 150 μM (final DMSO concentration 1%). The protein concentration was determined using the reduced heme Soret peak at 433 nm (149,000 M^{-1} cm^{-1}).[40] Protein concentration was adjusted after desalting the excess NO by comparing the A_{280} peaks to the unliganded protein. Activity assays were conducted at 25 °C and pH 7.5 in Buffer D, supplemented with 5 mM DTT and 5 mM $MgCl_2$. Reactions were initiated with 2 mM GTP and timepoints from were quenched with 125 mM $Zn(CH_3CO_2)_2$, followed by 125 mM $Na_2(CO_3)$ to adjust to pH 7. Samples were frozen at –80 °C until quantification. Quenched assays were thawed, and the zinc precipitate was spun down for 10 min at 23,150 *g*. The reactions were diluted by one to three orders of magnitude, and the cGMP was quantified in duplicate using an extracellular cGMP Enzyme Linked Immunosorbent Assay, following the manufacturer's instructions (Enzo Life Sciences). Concentrations of cGMP were determined from a standard curve, generated over 0.16–500 pmol/mL. Initial rates were calculated from the linear phase of the time course, where 5–10% of the GTP substrate was consumed. The experiment was repeated at least three times to ensure reproducibility.

Cryo-EM Sample Preparation and Data Collection

Samples were reduced in an anaerobic glove bag with 5 mM $Na_2S_2O_4$ for 15 minutes, and then desalted with a BioSpin6 column equilibrated with Buffer F (25 mM triethanolamine, pH 7.5, 25 mM NaCl, 5 mM DTT, 0.22 μM filtered). The inactive protein sample was diluted after thawing from a single frozen stock to ~2–4 μM in 1–3% trehalose. The sample (3.5 μL) was applied to glow discharged UltrAUfoil 2/2 200 mesh gold grids (Quantifoil) and plunge-frozen using a vitrobot Mark IV (Thermo Fischer). Blotting was performed under 100% humidity with zero blot force for 3–6 seconds. sGC was dispersed in vitreous ice which displayed clear contrast transfer function information (Figure A.3A). Active sample was prepared in a similar manner, but 500 μM NO (from DEA NONOate) and 150 μM YC-1 were added (Supplemental S4a-b).

Grids were screened and imaged on a Talos Arctica (Thermal Fischer) operated at 200 kV. Complete imaging conditions are described in Table A.1. Micrographs were collected at 36,000X nominal magnification on a K3 direct electron detector (Gatan) in super-resolution counted mode at 0.5685 Å/pix. Serial EM was used for automated image shift data collection of 2,841 and 9,330

movies for the inactive and active sample, respectively. Movies were taken in 100 ms frames, totaling an electron dose of 60 electrons per movie.

Cryo-EM Data Processing

Inactive sGC movies were drift-corrected, gain-corrected, and binned to 1.137Å/pix in 7x5 patches using MotionCor2.[41] Micrographs were CTF-corrected using CTFFIND4 and single particles were manually picked for initial 2D classification within Relion 3.0.[42–44] Class averages representing the full length complex (Supplemental S2c) were used for template picking with Relion autopicker, resulting in 675,956 particles. Particles were imported into Cryosparc2 and pruned using 2D classification and 3D *ab initio* classification.[45] Initial 2D classification yielded a large number of falsely picked background particles and particles which represented broken sGC dimers. We suspect the background particles are due to the size of the complex, while the broken particles likely stem from damage at the air-water interface (Figure A.3C). These poor class averages were removed from further processing steps. In total, 59,165 final particles were refined using non-uniform refinement with default parameters (Figure A.3B). Active sGC movies were binned to 2.274 Å/pix before template-free Laplacian picking for initial particle selection.[44] Micrographs were manually cleaned by visual inspection and FFT quality. Initial single particles were pruned using iterative 2D and 3D techniques (Supplemental S4b, c), as described above for the inactive state. Similarly, many of the initial picked particles were background or broken in the Active dataset (Supplemental S3c) Blurred density at low thresholds was seen in a predicted region for the alpha-HNOX but did not resolve during processing and was masked away during the final reconstruction. Of note, both Inactive and Active datasets underwent exhaustive processing schemes, including Bayesian polishing, multibody refinement and focused classification, none of which improved the resolution or showed evidence of multiple conformations in a single dataset.

Model Building

Homology models for each domain were built using *Phyre2* and correspond to the following PDB entries: α H-NOX (2O0C), α PAS (4GJ4), α CC (3HLS), α CAT (3UVJ), β H-NOX (2O0C), β PAS (4GJ4), β CC (3HLS), β CAT (2WZ1).[46] Domains (with side chains removed) were rigid-body docked into the inactive reconstruction using the *fit_in_map* function of Chimera.[47] Clear density for heme in the β H-NOX domain enabled clear distinction of the two H-NOX domains. Linkers between the H-NOX and PAS domains are missing in the density, likely due to flexibility, and were left out of the model (α 238–278 and β184–206). Initial placement of the CAT dimer was based on a significant extension in the sequence of the α CAT C-terminus, which is visible as unmodeled density. Continuous density from the CAT dimer enabled assignment of the CC and PAS domains. The linker region between the PAS domains and the CC were manually modeled in COOT based on secondary structure predictions from PSIPRED.[48,49] Refinement using iterative rounds of *phenix.real_space_refine* and inspection in COOT led to the final carbon back bone trace of the inactive state. Active state modeling was based on rigid-body fitting of domains based on the inactive state. Helical density was apparent throughout the structure (with the exception of the α H-NOX, as mentioned above). Linker regions between domains were corrected with COOT and *phenix.real_space_refine*. Final model idealization was carried out using *phenix.model*_idealization and validated using Molprobity.[50]

Small-angle x-ray scattering in-line with size-exclusion chromatography (SEC-SAXS)

Samples were reduced in an anaerobic glove bag with 5 mM $Na_2S_2O_4$ for 15 minutes, concentrated to 60 µL at 50 µM (~7.5 mg/mL), and then three rounds of dialysis were performed in Buffer G (50 mM KH_2PO_4, pH 7.4, 150 mM NaCl, 2% glycerol). The activated sample was prepared with 500 µM NO (from DEA NONOate). The samples were sealed and run within 3 hours of being prepared; samples did not undergo freeze–thaw after the sample was prepared (the protein was freeze–thawed once after purification for initial storage).

In situ sample purification was accomplished through SEC to isolate well-folded proteins from aggregates and other impurities immediately before exposure to synchrotron X-ray radiation using a custom designed flow cell. SEC-SAXS was collected at the SIBYLS beamline (bl12.3.1) at the Advanced Light Source at Lawrence Berkeley National Laboratory, Berkeley, California.[51,52] X-ray wavelength was set at λ = 1.127 Å and the sample-to-detector distance was 2,105 mm, as determined by silver behenate calibration, resulting in scattering vectors, q, ranging from 0.01 $Å^{-1}$ to 0.4 $Å^{-1}$. The scattering vector is defined as $q = 4\pi\sin\theta/\lambda$, where 2θ is the scattering angle. Data was collected using a Dectris PILATUS3X 2M detector at 20 °C and processed as previously described.[53,54] Briefly, a SAXS flow cell was directly coupled with an Agilent 1260 Infinity HPLC system using a Shodex KW-803 column. The column was equilibrated with running Buffer E with a flow rate of 0.5 mL/min for inactive sGC and 0.55 mL/min for activated sGC. To achieve activation conditions, the buffer was continuously sparged with nitrogen gas and the column was equilibrated for at least 2 hours to maintain an anaerobic environment. Several NONOates with various half-lives were added to the running buffer to achieve activation of sGC. These NONOates include DEA NONOate and DETA NONOate, which spontaneously release NO with half-lives of 16 minutes and 56 hours, respectively. Each sample was run through the SEC-SAXS system and 3-second X-ray exposures were collected continuously over the 30-minute elution. The SAXS frames recorded prior to the protein elution peak were used to correct all other frames. The corrected frames were investigated by radius of gyration R_g derived by the Guinier approximation $I(q) = I(0)\exp(-q^2R_g^2/3)$ with the limits $q*R_g < 1.3$ (Figure A.5B). The elution peak was mapped by comparing the integral of ratios to background and R_g relative to the recorded frame using the program SCÅTTER (Figure A.5A).

The frames in the regions of least R_g variation were averaged and merged in SCÅTTER to produce the highest signal-to-noise SAXS curves. These merged SAXS curves were used to generate the Guinier plots, volumes-of-correlation (V_c), pair distribution functions, P(r), and normalized Kratky plots. The Guinier plot indicated an aggregation-free state of the protein (Figure A.5B). The P(r) function was used to determine the maximal dimension of the macromolecule (D_{max}) and estimate inter-domain distances (Fig. 4A).[55] P(r) functions were normalized based on the molecular weight (MW) of the assemblies, as determined by the calculated V_c.[56]

Eluent was subsequently split (4 to 1) between the SAXS line and a multiple wavelength detector (UV-vis), set to 432 and 280 nm, multi-angle light scattering (MALS), and refractometer. The ratios of the protein (280 nm) and Soret band (432 nm) of the heme from SEC were used to evaluate the ligation state of *Ms* sGC upon NO binding (Figure A.5A). MALS experiments were performed using an 18-angle DAWN HELEOS II light scattering detector connected in tandem to an Optilab refractive index concentration detector (Wyatt Technology). System normalization and calibration was performed with bovine serum albumin using a 60 µL sample at 10 mg/mL in the same SEC running buffer and a dn/dc value of 0.185 and 0.15 mL/g for inactive and active sGC respectively.

The MALS data was used to compliment the MWs calculated by the SAXS analysis and, being furthest downstream, the MALS peaks were used to align the SAXS and UV-vis peaks along the x-axis (elution volume in mL/min) to compensate for fluctuations in timing and band broadening (Figure A.5A). UV–vis data was integrated using Agilent Chemstation software and baseline corrected using Origin 9.1 (Figure A.5A). MALS and differential refractive index data was analyzed using Wyatt Astra 7 software to monitor the homogeneity of the sample molecular weights across the elution peak, complementing the SEC-SAXS signal validation (Figure A.5A).

Solution Structure Modeling

The cryo-EM refined crystal structure for inactive sGC was used to build an initial atomistic model; all missing loops and terminal residues were added using MODELLER.[57] A purpose-designed rigid body modeling pipeline was applied to the completed structure using BILBOMD to systematically explore conformational space for both the inactive and activated states of sGC (Video 4).[58] To obtain an inactive state model, the α H-NOX domain and both α terminal tails were moved, then all α and β domains were moved by allowing flexibility at the hinge region where the CCs meet the CAT domains, while holding the CAT domains fixed and still permitting flexibility of the terminal tails. This was then used to generate the 1-NO state model by moving the same regions as the inactive state in reverse order. Theoretical SAXS curves were produced by FOXS.[59,60] Each model generated through BILBOMD as compared to the experimental SAXS profiles to assess the goodness of fit. Multistate model ensembles for activated sGC were determined using MultiFOXS.[61]

Results

Characterization and activity of full-length Manduca sexta sGC

We selected *Manduca sexta* (*Ms*) sGC as a biochemically tractable protein which has 36% and 59% sequence identity when compared to the human homolog for the α and β monomers, respectively. A schematic of the domain organization is shown in Figure 2.1A. Additionally, the regulatory properties in response to NO and YC-1 mirror those of its human counterpart and YC-1 had been shown to increase the *Ms* sGC 1-NO state to maximal activity.[40] While full-length *Ms* sGC has only been previously reported to be expressed in *E. coli*, we used a baculoviral expression protocol to produce reliable quantities of stable, full-length protein (Figure A.1).[40] UV–visible absorption spectra of *Ms* sGC without ligands (unliganded, U), with xsNO, and with carbon monoxide (CO) show the expected ligand-dependent shifts of the Soret and Q bands, indicating the β H-NOX domain is competent to bind diatomic gases (Figure 2.1B).

We confirmed that *Ms* sGC displays catalytic activity similar to other well-characterized mammalian homologues, including *Homo sapiens* sGC. *Ms* sGC exhibits a low, basal activity in the unliganded state (71 ± 36 nmol/min/mg) and partial activation in the 1-NO state (309 ± 77 nmol/min/mg) (Figure 2.1C, Figure A.1C), consistent with previous reports of mammalian sGCs.[22] Maximal activity could be achieved by adding excess NO to the 1-NO sample (1,988 ± 131 nmol/min/mg), or by the addition the sGC stimulator YC-1 to the 1-NO state (1,522 ± 38 nmol/min/mg) (Figure 2.1C, Figure A.1D, Figure A.2). Taken together, *Ms* sGC displays biochemical properties comparable to well-characterized mammalian sGCs.

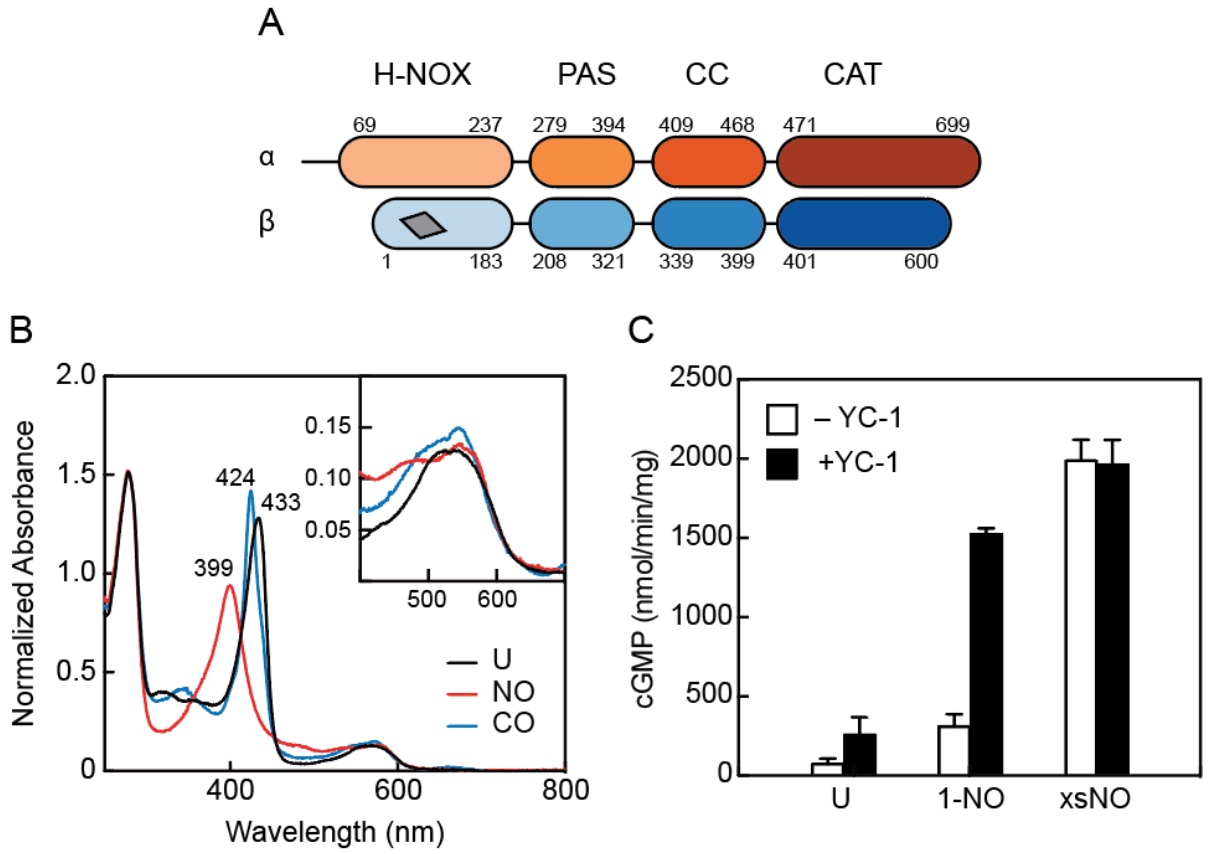


Figure 2.1 – Domain arrangement and activation of *Manduca sexta* (*Ms*) soluble guanylyl cyclase (sGC). A) Schematic representation of the *Ms* sGC heterodimer domain architecture. sGCs contain four domains: a heme nitric oxide/oxygen (H-NOX) binding domain, a Per/Arnt/Sim (PAS)-like domain, a coiled-coil (CC) domain, and a catalytic cyclase domain (CAT). The heme binding site in β H-NOX is represented by the grey quadrilateral. B) UV–visible absorption spectra of *Ms* sGC in the unliganded (U), NO-bound, and CO-bound states. The wavelength maxima of the Soret peaks are indicated. Inset: Q bands show increased splitting upon gas binding. C) Discontinuous cGMP activity assay for *Ms* sGC with various activation conditions: 1-NO, xsNO, and YC-1 ligands. Initial rates were taken from assays run at 25 °C, pH 7.5 with 2 mM Mg•GTP as the substrate (see Figure A.1D). cGMP formation was measured using an enzyme linked immunosorbent assay. The average initial rate is plotted, and the error bars reflect one standard deviation (n = 4).

Inactive conformation of sGC exhibits bent coiled-coils

Cryo-EM was used to solve the structure of full-length inactive *Ms* sGC. Two-dimensional classification after removing poor particles showed intact density with two lobes with connective density between them similar to previously reported EM envelopes obtained by negative stain (Figure A.3A-C).[19] Three-dimensional classification and refinement yielded a single reconstruction with a 5.1 Å nominal resolution (Figure A.3D) that exhibited a clear dimer interface and distinct CCs (Figure 2.2A, Table A.1). While particles display a slight orientation preference, local resolution is uniform throughout the density map (Figure 2.2E,F). Additionally, the overall map exhibits well-defined helices and continuous density. The two lobes of the structure are termed the "regulatory" lobe and the "catalytic" lobe, with the CC domains acting as a bridge between them (Figure 2.2A). The long axis of the regulatory lobe is positioned perpendicular to

the catalytic lobe and is predicted to contain the H-NOX/PAS bundle, while the CAT domains are predicted to form the catalytic lobe.[19]

Several features enabled conclusive assignment of the α and β subunit domains in the density map, despite the intermediate resolution and pseudosymmetry between the two subunits (Figure 2.2B). A rigid body global search of potential H-NOX positions revealed two possible positions for H-NOX domains within the regulatory lobe.[47] The αF helix of the β H-NOX domain (Figure A.1E) binds a heme cofactor while the α H-NOX does not. Only the position closest to the catalytic lobe contains density for the heme cofactor (Figure 2.2C), hence this was assigned to the β subunit. The two CAT domains could be differentiated because the α CAT domain contains a C-terminal extension (residues 661–699) which is absent in the β CAT domain. Only one of the domains in the catalytic lobe contains extra C-terminal density and thus the CAT domains could be assigned unambiguously (Figure A.3F). Connective density from the CAT domains enabled clear differentiation of the two CC and PAS domains (Figure 2.2D, green). As a result, the α PAS is assigned as the top domain, while β is the bottom domain (Figure 2.2B). Homology models for *Ms* sGC domains were combined representing residues α 51-250, 279-699 and β1-183, 205-597 to create a full-length model (Described in Methods).

The CC domains are in a parallel orientation, and both domains have a clear bend, distinct from a previously determined CC X-ray structure.[29] A linker extends from the C-terminus of the PAS domains and forms a rigid loop connecting the bent helix to the H-NOX/PAS domains (Figure 2.2D, light and dark green). The CC density bends sharply at residues α A422 and β L343. This bent region forms a buried helix, and along with the C-terminal PAS linker, creates an interaction nexus between the H-NOX, PAS, and CC domains (Figure 2.2D, highlighted in light and dark purple). The bent N-terminal portion of the CC domains (α 403–415 and β 333–345), is highly conserved across sGC homologues (Figure A.3I).

The CAT dimer displays a wreath-like fold with the monomers related by a twofold axis, typical for class III nucleotide cyclase domains.[20,62] The CAT domains align well with a previous structure of an inactive guanylate cyclase (Cα RMSD of 1.3 Å to PDB ID: 4NI2), displaying an inaccessible nucleotide-binding pocket that would require a significant rearrangement for activation (Figure A.3H). In the absence of a substrate-bound structure of a guanylate cyclase, we compared our model to a substrate-bound adenylate cyclase. The alignment of the α chain of an active adenylate cyclase structure (PDB ID: 1CJK, gray) with the α CAT domain of our inactive *Ms* sGC structure reveals that the Cα of β N538 from the β CAT domain is within 2 Å of the bound nucleotide analog in the adenylate cyclase structure (Figure 2.2E). This state is thus in a closed conformation that is sterically incompatible with nucleotide binding, explaining the lack of guanylate cyclase activity.

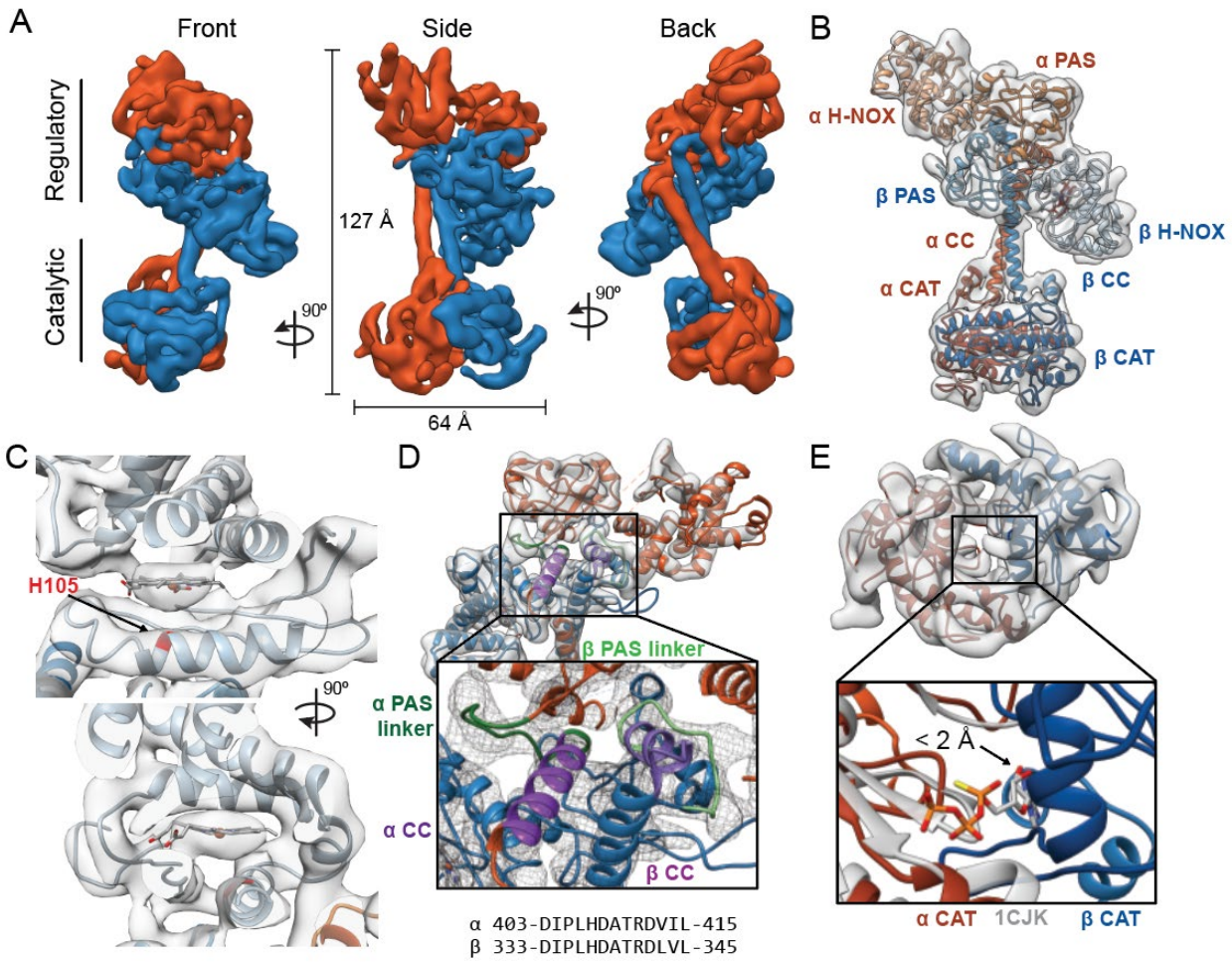


Figure 2.2 – Inactive *Ms* sGC forms a bent coiled-coil structure. A) Views of cryo-EM density for the inactive state, colored by subunit (α - orange, β - blue). Dimensions of the complex are shown in black. The two lobes of the enzyme are denoted "regulatory" and "catalytic". B) Molecular model of *Ms* sGC with domains labeled and colored as in Figure 2.1A. The heme is colored gray. C) Two views of the β H-NOX domain with heme shown in gray and H105 in red D) View of the bent coiled-coil (purple) and PAS linker (green) shown in shades for the α (dark) and β (light) dimer. Close up shows connective density between the PAS and CC domains (threshold 12σ). The sequences of the bent CC domains are shown below. E) View of the inactive catalytic dimer fit into density (α - orange, β - blue). Close up shows aligned active adenylyl cyclase (α CAT domain) and nucleotide in gray (PDB: 1CJK) compared with the inactive sGC model. The distance between the substrate analogue and nearest backbone Cα (β N538) is shown.

Activated sGC extends the regulatory lobe from catalytic core

To elucidate conformational changes associated with sGC activation, a structure of sGC bound to NO and the small molecule stimulator YC-1 was determined. We elected to supplement the xsNO *Ms* sGC with YC-1 to generate the most stable activated conformation. Particles of active *Ms* sGC were well-dispersed in vitreous ice (Figure A.4A,4B). Cleaned two-dimensional class averages exhibit a two-lobed density that is distinct from the inactive structure (Figure 2.3C). The final 3D reconstruction of the active state at 5.8 Å resolution displays clear density for the CCs, β H-NOX, and α helices throughout the structure (Figure 2.3A, Figure A.4D-F). Only diffuse density

was present for the α H-NOX, likely due to increased flexibility of this domain and it was thus omitted in the model (Figure 2.3B).

The orientation of the domains in the regulatory lobe is maintained from the inactive state to the active state, and the β H-NOX position is readily identified via the bound heme cofactor density (Figure 2.3C). A triangular-shaped lobe of unassigned density is present in the active state map but not in the inactive state map (Figure 2.3C, arrow). This density is located directly between the β H-NOX domain and the α and β CC domains, adjacent to either the A or B pyrrole of the heme (opposite of the propionate chains). The volume likely corresponds to YC-1 (304.3 Da) and with the current resolution two orientations of this small molecule are possible (Figure 2.3C). Residues within 5 Å of the density include α CC 422-427 (AQDGLR), and β V39, F77, C78, Y112, and Q349.

Although the density for the β H-NOX and PAS domains is similar between the inactive and active states, the overall shape of the molecule is more linear (Figure 2.3B). Continuous, linear helices from α L406–L457 and β A333–Y386 are seen in the active state (Figure 2.3D, highlighted in purple). This conformation is more aligned with predicted CC length and is more similar to the previously published crystal structures in terms of the range of residues seen in an unbent coiled-coil [29,34]

The CAT heterodimer in the active state still forms a wreath-like geometry, but with the two monomers moved apart from one another roughly perpendicular to the CC domain axis (Figure 2.3E). An alignment of the α chain of an active adenylate cyclase structure (PDB ID: 1CJK, gray) with the α CAT domain of our active *Ms* sGC structure shows the Cα of β N538 is now greater than 4 Å from the bound nucleotide analog, providing sufficient space for the nucleotide to bind in the active site. The *Ms* sGC CAT domain thus adopts an open conformation with the apparent capacity to bind substrate.

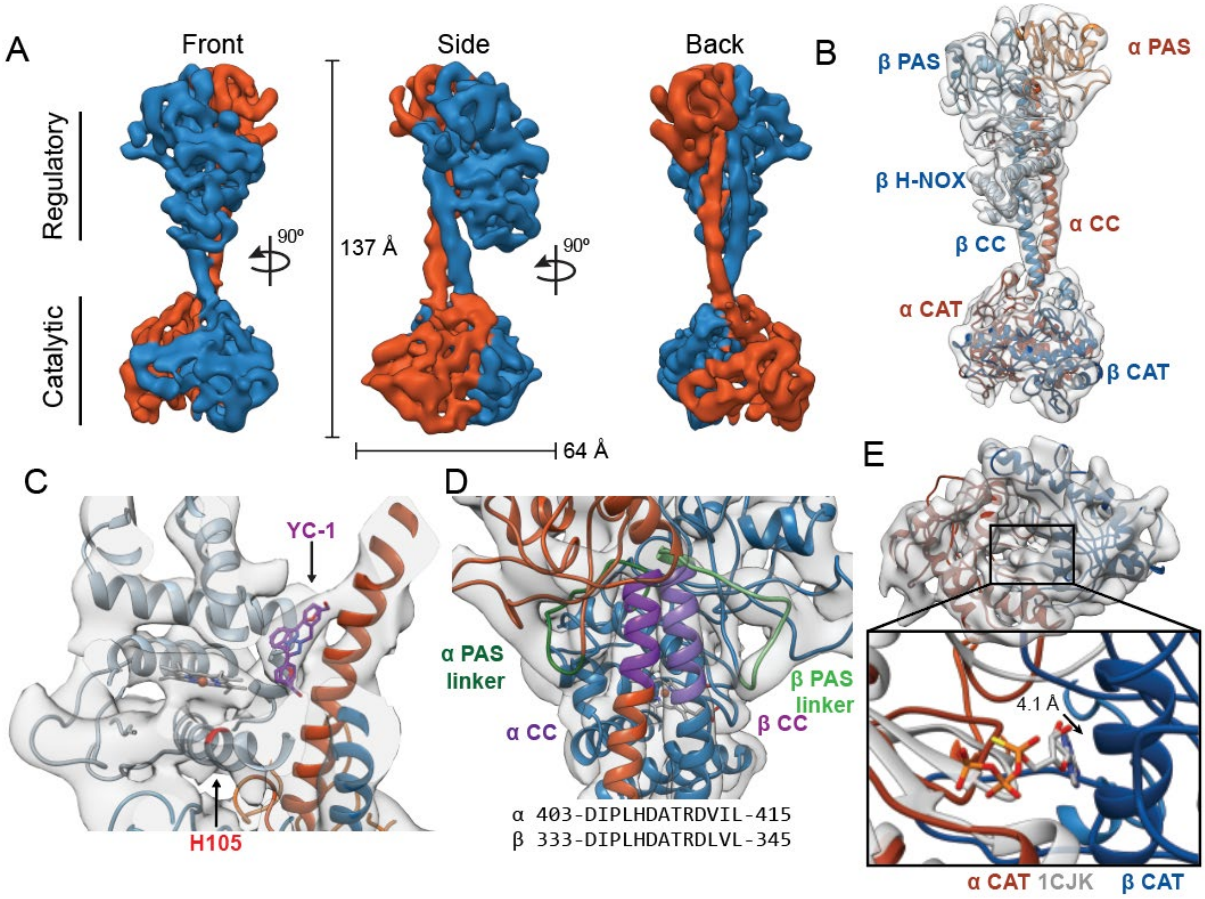


Figure 2.3 – Active *Ms* sGC forms an elongated structure. A) Views of cryo-EM density for the active state, colored by subunit (α - orange, β - blue). Dimensions of the complex are shown in black. B) Molecular model of *Ms* sGC colored as in Figure 2.1A. The heme is colored gray. The α H-NOX domain is not shown due to lack of density. C) View of the β H-NOX domain with heme shown in gray and β H105 in red. Two fits for the stimulator, YC-1, are shown in purple D) Model of the extended coiled-coil region (purple), highlighting the PAS linker region (green) shown in shades for the α (dark) and β (light) dime. The sequences of the bent CC region are shown below. E) View of active catalytic dimer (α - orange, β - blue) shown fit into the reconstruction density. Close up shows aligned active adenylate cyclase structure (PDB: 1CJK) with bound nucleotide (gray). The distance between the substrate analogue and Cα of the β N538 is shown.

Allosteric conformational rearrangement of sGC

Binding of NO to β H-NOX initiates the signal transduction event, thus interfaces were examined between the inactive and active states. A significant rearrangement of the quaternary structure of *Ms* sGC occurs upon activation with the regulatory domain rotating 71° and the catalytic domain rotating 40° (Figure 2.4A). In both inactive and active states, the β H-NOX domain is the closest domain of the regulatory lobe to the CAT domains (Figure 2.4A); however, in neither model does the β H-NOX domain come into direct contact with the CAT domains. In the inactive state, the Cα of β I47 is 12 Å away from the Cα of β H399 (Figure 2.4A, gray), too far for direct contact. The closest Cα-Cα distance of residues between the β H-NOX domain and the CAT domains in the active state is between β M33 and α W468 at 21Å (Figure 2.4A, black), remaining too far for direct contact between the β H-NOX and the CAT dimer. Therefore, allosteric

communication is transmitted through changes in CC positioning, not through direct H-NOX:CAT interactions.

Interaction surfaces, defined by a ≤ 10 Å Cα to Cα distance[63], occur between the β H-NOX and the PAS or CC region in both inactive and active states (Figure 2.4B,4C). In the inactive state, the αE, αF, and β2 secondary structural elements of the β H-NOX (82-126, green) form an extensive interface with the β PAS (270-275, red) and the α PAS-CC linker (400-419, blue) regions (Figure 2.4B, Figure A.1E). Overall, these interfaces are maintained in the active state (Figure 2.4C), in spite of the straightening by the CC domains. Activation of sGC induces the formation of an interface between β CC (355-367) and β H-NOX (33-41) (Figure 2.4C, purple). These residues on the β H-NOX comprise a loop located between helix αB and αC in the β H-NOX domain which was extended away from the CC in the inactive state. The large conformational change of the H-NOX/PAS bundle is necessary for the formation of the new interface in the active site, and thus this interaction could be critical for stabilizing the active sGC conformation.

In the active structure the CCs undergo a significant conformational change and rotation relative to the CAT dimer (Figure 2.4D). The active state CC domains are completely extended (56 Å and 63 Å for α and β respectively in the inactive state, to 74 Å for both CC domains the active state), as the bend present in the inactive state moves away from the CAT domains to be in line with the C-terminal portion of both CC domains. The CC conformational change upon activation twists the CC interface such that the β CC helix rotates relative to the α CC by 72 degrees (Figure 2.4D). The dramatic rearrangements of the H-NOX/PAS bundle leads to unbending of the CC domains, a change in their orientations as they project from the regulatory lobe, and finally to opening of the nucleotide binding pocket (Figure 2.4E). Alignment to the α CAT domain shows a clear rotation of the β CAT domain away from α CAT domain by 40° (Figure 2.4E, left). This rotation entails the pinching together of the base of the CC domains (Figure 2.4E, middle). In combination, these motions open the GTP binding pocket by moving the β CAT active site helix (535-545) away from the α CAT domain and into an active conformation.

SAXS shows a distribution of inactive and active states in solution

Small angle X-ray scattering (SAXS) was used to interrogate conformational changes with and without NO in solution. Experiments with sGC stimulators were not included as the limited solubility of YC-1 is precludes the collection of scattering data. Size exclusion chromatography (SEC)-SAXS and SEC-multi angle light scattering (MALS) chromatograms show a symmetrical peak for both samples with little variation in radius of gyration (R_g) and estimated MW across the peaks, indicating sample homogeneity (Figure A.5A).

The UV-visible absorption of the Soret band of the heme and of the protein itself were used to evaluate the ligation state of *Ms* sGC in the SAXS experiment. The ratios of Soret band to protein peak in the full UV-visible absorption spectra are 0.85, 0.35, and 0.20 in the unliganded, 1-NO, and xsNO states, respectively (Figure 2.1B, Figure A.5A). The ratios of these two peaks in the SEC-SAXS experiments were 0.77 with NO and 0.32 without NO (Figure A.5A). The 432/280 ratio in the SAXS sample without NO is less than in the full UV-Visible spectra, indicating that the sample had a slightly lower heme incorporation. The ratio for the SAXS sample with NO was lower than that expected for 1-NO, but much larger than predicted for the xsNO ratio, indicating that the SAXS samples were in the unliganded and 1-NO state during elution, respectively.

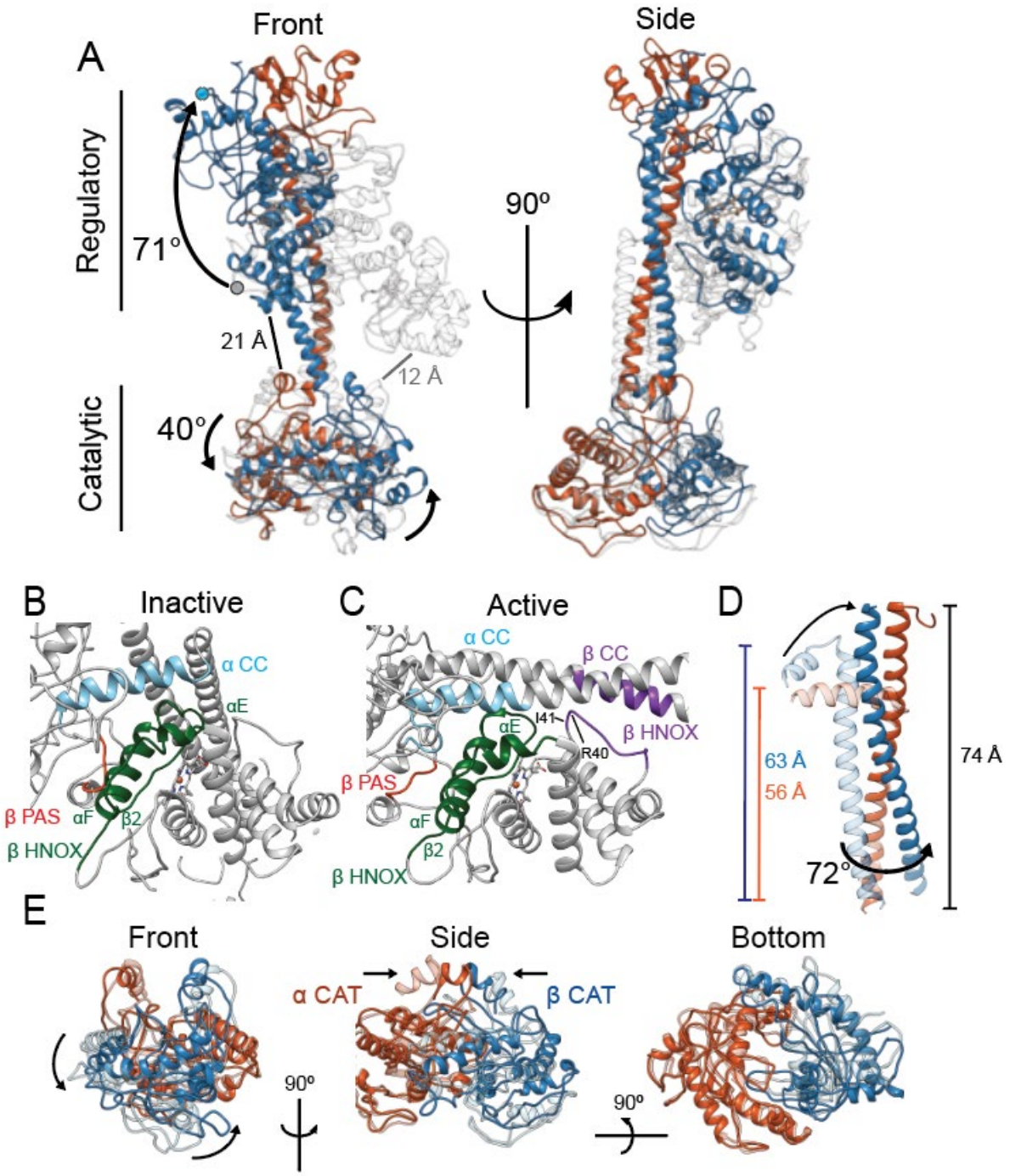


Figure 2.4 – Conformational rearrangements of *Ms* sGC upon activation. A) Overlay of the inactive (transparent, gray) and active (α - orange, β – blue) shown in two views. Rotation of the regulatory domain and CAT dimer are shown with arrows and label with degree of rotation. Distances between the β H-NOX domain and CAT dimer are labeled for inactive (gray) and active (black). B) Close up view of the inactive interfaces between the β HNOX (green), β PAS (red), and α CC (blue). C) Close up view of the active interfaces colored as in Figure 2.4B with the β H-NOX:β CC interface in purple. D) Overlay of the inactive (transparent) and active (colored) CC domains when aligned to the active α CC domain. Dimensions of the CC and rotation are labeled in color for the inactive and black for the active. E) Three views of aligned inactive (transparent) and active (colored) CAT dimers are shown. Direction of rotation (curved arrows) and pinching (arrows) are shown.

The R_g for the unliganded and 1-NO state were 43.1 ± 0.4 and 43.8 ± 0.2 Å, respectively, representing an elongation of the structure (Figure A.5A, Table A.2). Additionally, the Pair-distance function (P(r)), which shows a composite of the inter-atomic distances, displays a distinct shift between the unliganded and 1-NO states in the region from ~70 to ~100 Å. This corresponds to an extension of the regulatory lobe from catalytic lobe (Figure 2.5A). Shifts in the secondary peaks of the Kratky plots (Figure A.5B) further suggest separation of the regulatory and catalytic lobes upon activation.

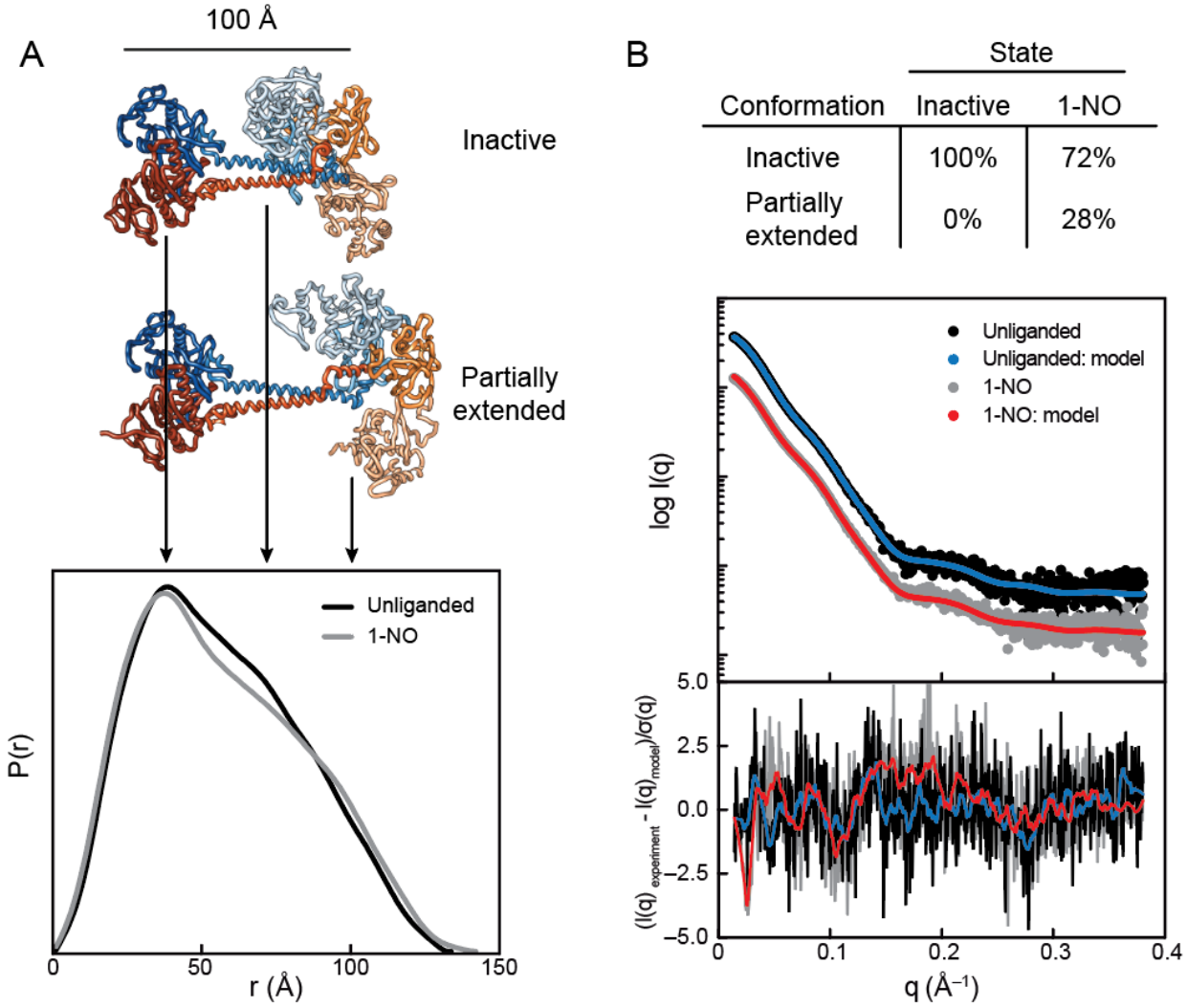


Figure 2.5 – SAXS analysis of the activation of sGC. A) Normalized P(r) for the unliganded (black) and 1-NO (gray) states of sGC are shown together with corresponding models of the inactive and partially extended conformations of sGC (N-terminal and C-terminal tails are omitted for clarity). The area of each P(r) is normalized relative to the SAXS-determined molecular weights (Table A.2). B) Experimental (black and gray) and theoretical (blue and red) SAXS profiles for the solution state models in the panel a. SAXS fits are shown together with the fit residuals in the below graph.

Using the inactive model obtained from cryo-EM as an initial Cα model, a rigid-body modeling pipeline was developed to systematically explore conformational space (see Materials and Methods).[58] A minimal ensemble search was performed over thousands of sGC conformations

and corresponding scattering curves were calculated and compared to the experimental data. The result of this minimal ensemble search confirmed the presence of a single sGC conformation while in an inactive state (Figure 2.5B, Figure A.5C), which overlays with the inactive structure with an RMSD of 2.0 Å. The 1-NO state was best modeled as an ensemble of two states, with the majority (72%) of the sample consistent with the inactive model (Figure 2.5B). However, the remainder of the sample (28%) is consistent with a more elongated conformation, one that is in between the inactive and active state obtained by cryo-EM. The Cα RMSD of the partially active SAXS model and the active state cryo-EM model is 13.5 Å. This analysis shows that sGC adopts a mixture of inactive and active conformations in the 1-NO state, suggesting that NO binding at the heme establishes an equilibrium between these two conformational extremes (Figure A.5C).

Discussion

The structures of the inactive and active states of full-length sGC provide detailed insight into the molecular mechanism of activation of *Ms* sGC. The most unexpected feature of the inactive state of the enzyme are the bends in the CC domains. Straightening of the bent CCs is the clearest structural link between the regulatory and catalytic lobes during activation. Previous HDX-MS data showed a differential exchange pattern for the CC domains, where the N-terminal portion of the α CC and the C-terminal portion of the β CC increase in H/D exchange upon NO binding, while the C-terminal portion of the α CC and the N-terminal portion of the β CC decrease in H/D exchange (Figure A.4G)[36]. The changes seen in HDX in the CCs near the bend are thus consistent with the large-scale rearrangements of the CCs that we have now observed in the cryo-EM. Furthermore, the bent portion of the CC domain is highly conserved among sGC homologues, consistent with its centrality to allosteric communication.

The C-terminal portions of the CC domains undergo a 72° twist and contain a motif known as the signaling helix (or S-helix)[64]. Similar to sGC, proteins with this motif contain both receptor and output domains that are connected by dimeric coiled-coils. The inactive and active sGC structures are the first full-length enzyme with a S-helix motif to be characterized. This motif aligns with C-terminal part of the CC domain, part of the 72° twist described above (Figure A.3J). It is possible that the twisting motion of these domains is a more general mechanism of activation for proteins with S-helices.

In transitioning between the inactive and active states, the H-NOX/PAS bundle rotates 71° in order for β H-NOX residues 33-41 to form an interface with residues β 355-367 of the β CC domain (Figure 2.4C, purple). The αH helix of the β H-NOX domain contains the proximal histidine that binds to the heme cofactor.[18] Crystal structures of bacterial H-NOX domains with and without NO show a rotation in the αF helix of the β H-NOX. However, in our full-length structures, the αF of the β H-NOX retains its interface with the β PAS and the α CC domain in both states (Figure 2.4B,4C). More strikingly, binding of YC-1 at a site bridging the β H-NOX and CC domain highlights how stabilization of this contact drives CC unbending (Figure 2.3C). This interface corroborates specific point mutations known to decrease activity of the full-length protein; specifically, single point mutants of β I41E in the β H-NOX domain retain basal activity but are only minimally activate with excess NO, implying that these sGC variants are catalytically competent but not as sensitive to stimulation as the wildtype enzyme.[39,65] The structural finding that residue Ile41 of the β H-NOX form contacts with the unbent β CC that are specific for the

active conformation explains these phenotypes. Destabilizing these contacts by mutating Ile41 thus blocks this interface and prevents sGC from reaching the active conformation. While this paper was under review, a cryo-EM map of activated *H. sapiens* sGC was reported, where the activation was achieved by using xsNO.[66] This xsNO map and the xsNO + YC-1 map reported here overlay well, except for the proposed density for YC-1. This lends support to our assignment of the YC-1 binding site at the β H-NOX- β CC interface. These insights suggest that the β H-NOX- β CC contact may be the most critical allosteric switch in the regulatory lobe of sGC.

While conformational changes are readily observed by comparing the two cryo-EM structures, the SAXS implies that the 1-NO state can sample similar conformations (Figure 2.5A). Although only a fraction of the population exhibits an extension, inspection of the Soret to protein UV-visible absorption indicates that the protein is in the 1-NO state. These data support the hypothesis that when the first NO binds to the heme of the β H-NOX, sGC adopts an equilibrium between the inactive and a partially extended conformation, with a K_{eq} = [active]/[inactive] = [0.72]/[0.28] = 0.39. This is corroborated by the activity data from the 1-NO state, which exhibits 15% of the maximal activity. The conformational heterogeneity observed in SAXS analysis could represent the physiological state of sGC at basal cellular conditions.

In the absence of increased NO concentrations, sGC stimulators can maximally activate cGMP production and are now used to treat forms of pulmonary hypertension.[67,68] YC-1 was present in the sample used to generate the active state reconstruction of *Ms* sGC, and extra density is seen near the new β H-NOX : β CC interface (Figure 2.3C). Distinct changes in both resonance Raman and electron paramagnetic resonance spectroscopy signatures for heme ligands have been detected with both YC-1 and BAY 41-2272 supporting this proposed binding site.[69] Previously, a stimulator binding site was proposed between helices αA and αD in the β H-NOX domain based on cross linking and NMR data.[70] However, there is no density for YC-1 present in the active state reconstruction between αA and αD helices. We note that our study used YC-1 compared to IWP-854, IWP-051 and BAY 41-2272 previously used, which could explain the difference in binding sites. Visualization of YC-1 bound to sGC is an important development as a proof of concept that small molecule sGC stimulators can now be characterized structurally.

Adenylyl cyclases (AC) catalyze a similar reaction mechanism as sGC. Mammalian membrane-associated ACs contain two catalytic domains which form the intrapolypeptide equivalent of a heterodimer. The catalytic domains in active mammalian AC heterodimer rotate by 7° with respect to an inactive homodimeric counterpart bound to two molecules of forskolin[62,71,72]. We observe a rotation of the sGC CAT domains about the same axis, although the magnitude is larger for sGC, at 40°. A full-length membrane bound AC (AC-9) was recently solved in the active nucleotide bound state.[34] The CC domains between the full-length active sGC and AC9 structures overlay well, consistent with a common active geometry for GCs and ACs. The AC9 has a very different angle between the CC and CAT domains. The rotation of the sGC regulatory lobe is sterically incompatible with the presence of a membrane, thus by their different natures as soluble and membrane-associated enzymes, sGC and ACs must differ in the detailed modes of allosteric communication.

To date a detailed understanding of sGC activation has been hampered by the lack of full-length structures. The inactive and active structures in tandem with established studies lead to

formation of new hypotheses for the activation of sGC. First, activity assays with sGC truncations suggested that direct interaction of the β H-NOX domain and the CAT dimer was responsible for sGC inhibition.[37] However, the structural data shows no direct interaction between the β H-NOX and the CAT domain in either conformation. Instead, the formation of the new β H-NOX/β CC interface in the active state stimulates the active CAT conformation. Activation leads to a global conformational rearrangement of the heterodimer, elongating the structure; however, only a single equivalent of NO is required to cleave the bond between the heme Fe center and the ligating histidine residue. Maximal activation of sGC involves either the binding of a second NO molecule to a non-heme site on sGC or a small molecule stimulator stabilizing the active state.[73,74]

Figure 2.6 depicts the proposed physiological activation sequence of sGC, where unliganded sGC adopts a more compact conformation with bent CC domains. Upon NO binding to the heme, an equilibrium of conformational states is established, with the partially elongated state affording about 15% of the maximal activity. Given the very tight association between NO and the heme, this represents the basal cellular activity. Finally, in the presence of either an increase in NO concentration or sGC stimulating molecules, the completely extended, fully active state is reached and sGC reaches maximal catalytic activity. Using these structures as a starting point, new avenues of exploration can now be undertaken, to elucidate the molecular mechanisms of excess NO activation. Critical residues for interdomain interactions along with the proposed stimulator binding site can be characterized in more detail. Having resolved the two structural states of an important therapeutic target, the structures of sGC will influence rational design of improved drugs for diseases associated with NO signaling impairment.

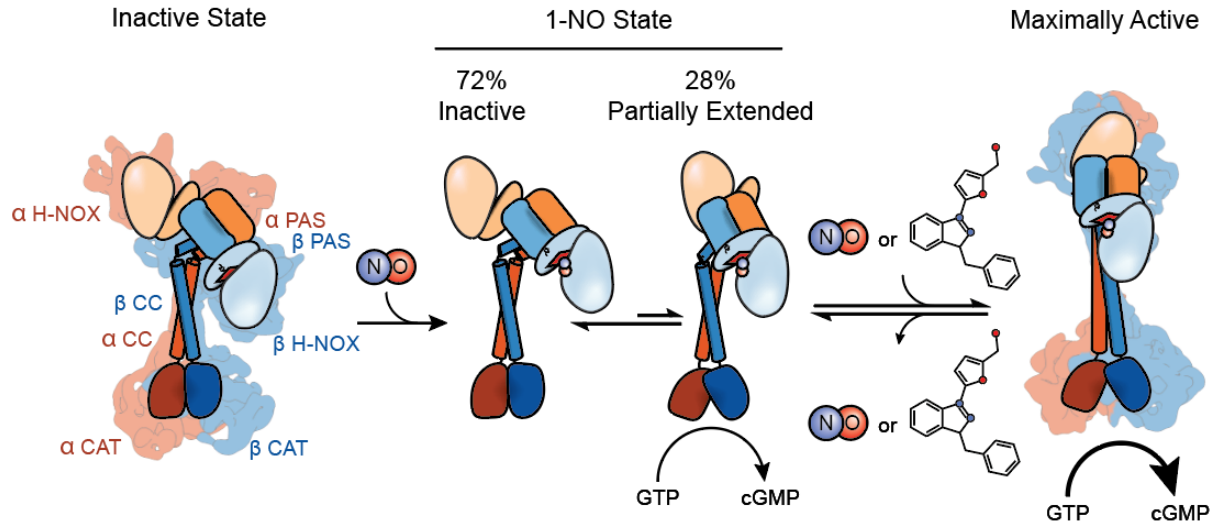


Figure 2.6 – Model for conformational rearrangement upon sGC activation. Schematic of the sGC activation pathway. sGC adopts a bent CC conformation after formation of the holoenzyme. Upon the addition of NO an equilibrium exists between the inactive and partially extended states, conferring partial activation of sGC. When excess NO enters the cell or the addition of a stimulator compound, the equilibrium shifts to the active state, allowing for an open CAT dimer conformation and maximal activity.

References

(1) Horst, B. G.; Yokom, A. L.; Rosenberg, D. J.; Morris, K. L.; Hammel, M.; Hurley, J. H.; Marletta, M. A. Allosteric Activation of the Nitric Oxide Receptor Soluble Guanylate Cyclase Mapped by Cryo-Electron Microscopy. *Elife* **2019**, *8*, e50634.

(2) Palmer, R. M.; Ferrige, A. G.; Moncada, S. Nitric Oxide Release Accounts for the Biological Activity of Endothelium-Derived Relaxing Factor. *Nature* **1987**, *327* (6122), 524–526.

(3) Moncada, S.; Palmer, R. M.; Higgs, E. A. Nitric Oxide: Physiology, Pathophysiology, and Pharmacology. *Pharmacol. Rev.* **1991**, *43* (2), 109–142.

(4) Arnold, W. P.; Mittal, C. K.; Katsuki, S.; Murad, F. Nitric Oxide Activates Guanylate Cyclase and Increases Guanosine 3':5'-Cyclic Monophosphate Levels in Various Tissue Preparations. *Proc. Natl. Acad. Sci. U. S. A.* **1977**, *74* (8), 3203–3207.

(5) Stone, J. R.; Marletta, M. A. Soluble Guanylate Cyclase from Bovine Lung: Activation with Nitric Oxide and Carbon Monoxide and Spectral Characterization of the Ferrous and Ferric States. *Biochemistry* **1994**, *33* (18), 5636–5640.

(6) Russwurm, M.; Koesling, D. NO Activation of Guanylyl Cyclase. *EMBO J.* **2004**, *23* (22), 4443–4450.

(7) Derbyshire, E. R.; Marletta, M. A. Structure and Regulation of Soluble Guanylate Cyclase. *Annu. Rev. Biochem.* **2012**, *81* (1), 533–559.

(8) Montfort, W. R.; Wales, J.; Weichsel, A. Structure and Activation of Soluble Guanylyl Cyclase, the Nitric Oxide Sensor. *Antioxidants redox Signal.* **2016**, *00* (00), 1–16.

(9) Horst, B. G.; Marletta, M. A. Physiological Activation and Deactivation of Soluble Guanylate Cyclase. *Nitric Oxide* **2018**, *77*, 65–74.

(10) Lucas, K. A.; Pitari, G. M.; Kazerounian, S.; Ruiz-Stewart, I.; Park, J.; Schulz, S.; Chepenik, K. P.; Waldman, S. a. Guanylyl Cyclases and Signaling by Cyclic GMP. *Pharmacol. Rev.* **2000**, *52* (3), 375–414.

(11) Prast, H.; Philippu, A. Nitric Oxide as Modulator of Synaptic Function. *Prog. Neurobiol.* **2001**, *64*, 51–68.

(12) Benarroch, E. E. Nitric Oxide: A Pleiotropic Signal in the Nervous System. *Neurology* **2011**, *77* (16), 1568–1576.

(13) Bredt, D. S. Endogenous Nitric Oxide Synthesis: Biological Functions and Pathophysiology. *Free Radic. Res.* **1999**, *31* (6), 577–596.

(14) Friebe, A.; Koesling, D. The Function of NO-Sensitive Guanylyl Cyclase: What We Can Learn from Genetic Mouse Models. *Nitric Oxide - Biology and Chemistry*. 2009, pp 149–156.

(15) Lundberg, J. O.; Gladwin, M. T.; Weitzberg, E. Strategies to Increase Nitric Oxide Signalling in Cardiovascular Disease. *Nat. Rev. Drug Discov.* **2015**, *14* (9), 623–641.

(16) Hervé, D.; Philippi, A.; Belbouab, R.; Zerah, M.; Chabrier, S.; Collardeau-Frachon, S.; Bergametti, F.; Essongue, A.; Berrou, E.; Krivosic, V.; et al. Loss of A1β1 Soluble Guanylate Cyclase, the Major Nitric Oxide Receptor, Leads to Moyamoya and Achalasia. *Am. J. Hum. Genet.* **2014**, *94* (3), 385–394.

(17) Wallace, S.; Guo, D. C.; Regalado, E.; Mellor-Crummey, L.; Bamshad, M.; Nickerson, D. A.; Dauser, R.; Hanchard, N.; Marom, R.; Martin, E.; et al. Disrupted Nitric Oxide Signaling Due to GUCY1A3 Mutations Increases Risk for Moyamoya Disease, Achalasia and Hypertension. *Clin. Genet.* **2016**, *90* (4), 351–360.

(18) Zhao, Y.; Schelvis, J.; Babcock, G.; Marletta, M. A. Identification of Histidine 105 in the Beta1 Subunit of Soluble Guanylate Cyclase as the Heme Proximal Ligand. *Biochemistry* **1998**, *37* (97), 4502–4509.
(19) Campbell, M. G.; Underbakke, E. S.; Potter, C. S.; Carragher, B.; Marletta, M. A. Single-Particle EM Reveals the Higher-Order Domain Architecture of Soluble Guanylate Cyclase. *Proc. Natl. Acad. Sci.* **2014**, *111* (8), 2960–2965.
(20) Hurley, J. H. The Adenylyl and Guanylyl Cyclase Superfamily. *Curr. Opin. Struct. Biol.* **1998**, *8*, 770–777.
(21) Cary, S. P. L.; Winger, J. A.; Marletta, M. A. Tonic and Acute Nitric Oxide Signaling through Soluble Guanylate Cyclase Is Mediated by Nonheme Nitric Oxide, ATP, and GTP. *Proc. Natl. Acad. Sci.* **2005**, *102* (37), 13064–13069.
(22) Fernhoff, N. B.; Derbyshire, E. R.; Marletta, M. A. A Nitric Oxide/Cysteine Interaction Mediates the Activation of Soluble Guanylate Cyclase. *Proc. Natl. Acad. Sci.* **2009**, *106* (51), 21602–21607.
(23) Herzik, M. A.; Jonnalagadda, R.; Kuriyan, J.; Marletta, M. A. Structural Insights into the Role of Iron–Histidine Bond Cleavage in Nitric Oxide-Induced Activation of H-NOX Gas Sensor Proteins. *Proc. Natl. Acad. Sci.* **2014**, *111* (40), E4156–E4164.
(24) Zhao, Y.; Brandish, P.; Ballou, D.; Marletta, M. A. A Molecular Basis for Nitric Oxide Sensing by Soluble Guanylate Cyclase. *Proc. Natl. Acad. Sci. U. S. A.* **1999**, *96* (26), 14753–14758.
(25) Stasch, J. P.; Pacher, P.; Evgenov, O. V. Soluble Guanylate Cyclase as an Emerging Therapeutic Target in Cardiopulmonary Disease. *Circulation* **2011**, *123* (20), 2263–2273.
(26) Ko, F. N.; Wu, C. C.; Kuo, S. C.; Lee, F. Y.; Teng, C. M. YC-1, a Novel Activator of Platelet Guanylate Cyclase. *Blood* **1994**, *84* (12), 4226–4233.
(27) Pellicena, P.; Karow, D. S.; Boon, E. M.; Marletta, M. A.; Kuriyan, J. Crystal Structure of an Oxygen-Binding Heme Domain Related to Soluble Guanylate Cyclases. *Proc. Natl. Acad. Sci. U. S. A.* **2004**, *101*, 12854–12859.
(28) Purohit, R.; Weichsel, A.; Montfort, W. R. Crystal Structure of the Alpha Subunit PAS Domain from Soluble Guanylyl Cyclase. *Protein Sci.* **2013**, *22* (10), 1439–1444.
(29) Ma, X.; Beuve, A.; van den Akker, F. Crystal Structure of the Signaling Helix Coiled-Coil Domain of the Beta1 Subunit of the Soluble Guanylyl Cyclase. *BMC Struct. Biol.* **2010**, *10* (2).
(30) Winger, J. A.; Derbyshire, E. R.; Lamers, M. H.; Marletta, M. A.; Kuriyan, J. The Crystal Structure of the Catalytic Domain of a Eukaryotic Guanylate Cyclase. *BMC Struct. Biol.* **2008**, *8*, 42–52.
(31) Allerston, C. K.; von Delft, F.; Gileadi, O. Crystal Structures of the Catalytic Domain of Human Soluble Guanylate Cyclase. *PLoS One* **2013**, *8* (3), e57644.
(32) Seeger, F.; Quintyn, R.; Tanimoto, A.; Williams, G. J.; Tainer, J. A.; Wysocki, V. H.; Garcin, E. D. Interfacial Residues Promote an Optimal Alignment of the Catalytic Center in Human Soluble Guanylate Cyclase: Heterodimerization Is Required but Not Sufficient for Activity. *Biochemistry* **2014**, *53* (13), 2153–2165.
(33) Vercellino, I.; Rezabkova, L.; Olieric, V.; Polyhach, Y.; Weinert, T.; Kammerer, R. A.; Jeschke, G.; Korkhov, V. M. Role of the Nucleotidyl Cyclase Helical Domain in Catalytically Active Dimer Formation. *Proc. Natl. Acad. Sci.* **2017**, *114* (46), E9821–E9828.
(34) Qi, C.; Sorrentino, S.; Medalia, O.; Korkhov, V. M. The Structure of a Membrane

Adenylyl Cyclase Bound to an Activated Stimulatory G Protein. *Science (80-.).* **2019**, *364* (6438), 389–394.
(35) Fritz, B.; Roberts, S.; Ahmed, A.; Breci, L.; Li, W.; Weichsel, A.; Brailey, J.; Wysocki, V.; Tama, F.; Montfort, W. Molecular Model of a Soluble Guanylyl Cyclase Fragment Determined by Small-Angle X-Ray Scattering and Chemical Cross-Linking. *Biochemistry* **2013**, *52* (9), 1568–1582.
(36) Underbakke, E. S.; Iavarone, A. T.; Chalmers, M. J.; Pascal, B. D.; Novick, S.; Griffin, P. R.; Marletta, M. A. Nitric Oxide-Induced Conformational Changes in Soluble Guanylate Cyclase. *Cell Struct.* **2014**, *22* (4), 602–611.
(37) Winger, J. A.; Marletta, M. A. Expression and Characterization of the Catalytic Domains of Soluble Guanylate Cyclase: Interaction with the Heme Domain. *Biochemistry* **2005**, *44* (10), 4083–4090.
(38) Baskaran, P.; Heckler, E. J.; Van Den Akker, F.; Beuve, A. Aspartate 102 in the Heme Domain of Soluble Guanylyl Cyclase Has a Key Role in No Activation. *Biochemistry* **2011**, *50* (20), 4291–4297.
(39) Underbakke, E. S.; Iavarone, A. T.; Marletta, M. A. Higher-Order Interactions Bridge the Nitric Oxide Receptor and Catalytic Domains of Soluble Guanylate Cyclase. *Proc. Natl. Acad. Sci.* **2013**, *110* (17), 6777–6782.
(40) Hu, X.; Murata, L. B.; Weichsel, A.; Brailey, J. L.; Roberts, S. A.; Nighorn, A.; Montfort, W. R. Allostery in Recombinant Soluble Guanylyl Cyclase from Manduca Sexta. *J. Biol. Chem.* **2008**, *283* (30), 20968–20977.
(41) Zheng, S. Q.; Palovcak, E.; Armache, J.-P.; Verba, K. A.; Cheng, Y.; Agard, D. A. MotionCor2: Anisotropic Correction of Beam-Induced Motion for Improved Cryo-Electron Microscopy. *Nat. Methods* **2017**, *14* (4), 331–332.
(42) Rohou, A.; Grigorieff, N. CTFFIND4: Fast and Accurate Defocus Estimation from Electron Micrographs. *J. Struct. Biol.* **2015**, *192* (2), 216–221.
(43) Scheres, S. H. W. RELION: Implementation of a Bayesian Approach to Cryo-EM Structure Determination. *J. Struct. Biol.* **2012**, *180* (3), 519–530.
(44) Nakane, T.; Kimanius, D.; Lindahl, E.; Scheres, S. H. Characterisation of Molecular Motions in Cryo-EM Single-Particle Data by Multi-Body Refinement in RELION. *Elife* **2018**, *7*, 1–18.
(45) Punjani, A.; Rubinstein, J. L.; Fleet, D. J.; Brubaker, M. A. CryoSPARC: Algorithms for Rapid Unsupervised Cryo-EM Structure Determination. *Nat. Methods* **2017**, *14* (3), 290–296.
(46) Kelly, L. A.; Mezulis, S.; Yates, C. M.; Wass, M. N.; Sternberg, M. J. . The Phyre2 Web Portal for Protein Modelling, Prediction, and Analysis. *Nat. Protoc.* **2015**, *10* (6), 845–858.
(47) Pettersen, E. F.; Goddard, T. D.; Huang, C. C.; Couch, G. S.; Greenblatt, D. M.; Meng, E. C.; Ferrin, T. E. UCSF Chimera--a Visualization System for Exploratory Research and Analysis. *J. Comput. Chem.* **2004**, *25* (13), 1605–1612.
(48) Emsley, P.; Lohkamp, B.; Scott, W. G.; Cowtan, K. Features and Development of Coot . *Acta Crystallogr. Sect. D Biol. Crystallogr.* **2010**, *66* (4), 486–501.
(49) Vynne, J. The Application and Economic Benefits of Blasthole Drill Monitors. *Proc. ISEE 23rd Annu. Conf. Explos. Blasting Tech.* **1997**, 635–646.
(50) Chen, V. B.; Arendall, W. B.; Headd, J. J.; Keedy, D. A.; Immormino, R. M.; Kapral, G. J.; Murray, L. W.; Richardson, J. S.; Richardson, D. C. MolProbity : All-Atom Structure

Validation for Macromolecular Crystallography . *Acta Crystallogr. Sect. D Biol. Crystallogr.* **2010**, *66* (1), 12–21.
(51) Classen, S.; Rodic, I.; Holton, J.; Hura, G. L.; Hammel, M.; Tainer, J. A. Software for the High-Throughput Collection of SAXS Data Using an Enhanced Blu-Ice/DCS Control System. *J. Synchrotron Radiat.* **2010**, *17* (6), 774–781.
(52) Classen, S.; Hura, G. L.; Holton, J. M.; Rambo, R. P.; Rodic, I.; Mcguire, P. J.; Dyer, K.; Hammel, M.; Meigs, G.; Frankel, K. A.; et al. Implementation and Performance of SIBYLS : A Dual Endstation Small-Angle X-Ray Scattering and Macromolecular Crystallography Beamline at the Advanced Light Source. *J. Appl. Crystallogr.* **2013**, *46*, 1–13.
(53) Dyer, K. N.; Hammel, M.; Rambo, R. P.; Tsutakawa, S. E.; Rodic, I.; Classen, S.; Tainer, J. A.; Hura, G. L. *High-Throughput SAXS for the Characterization of Biomolecules in Solution : A Practical Approach*, Methods in.; Chen, Y. W., Ed.; Springer International Publishing, 2014; Vol. 1091.
(54) Hura, G. L.; Menon, A. L.; Hammel, M.; Rambo, R. P.; Ii, F. L. P.; Tsutakawa, S. E.; Jr, F. E. J.; Classen, S.; Frankel, K. A.; Hopkins, R. C.; et al. Robust , High-Throughput Solution Structural Analyses by Small Angle X-Ray Scattering (SAXS). *Nat. Methods* **2009**, *6* (8), 606–612.
(55) Putnam, C. D.; Hammel, M.; Hura, G. L.; Tainer, J. A. X-Ray Solution Scattering (SAXS) Combined with Crystallography and Computation: Defining Accurate Macromolecular Structures, Conformations and Assemblies in Solution. *Q. Rev. Biophys.* **2007**, *40* (3), 191–285.
(56) Rambo, R. P.; Tainer, J. A. Accurate Assessment of Mass , Models and Resolution by Small-Angle Scattering. *Nature* **2013**, *496* (7446), 477–481.
(57) Fiser, A.; Kinh, R.; Do, G.; Andrej, S. Modeling of Loops in Protein Structures. *Protein S* **2000**, *9*, 1753–1773.
(58) Pelikan, M.; Hura, G. L.; Hammel, M. Structure and Flexibility within Proteins as Identified through Small Angle X-Ray Scattering. *Gen. Physiol. Biophys.* **2009**, *28*, 174–189.
(59) Schneidman-duhovny, D.; Hammel, M.; Sali, A. FoXS : A Web Server for Rapid Computation and Fitting of SAXS Profiles. *Nucleic Acids Res.* **2010**, *38*, 540–544.
(60) Schneidman-duhovny, D.; Hammel, M.; Tainer, J. A.; Sali, A. Accurate SAXS Profile Computation and Its Assessment by Contrast Variation Experiments. *Biophys. J.* **2013**, *105* (4), 962–974.
(61) Schneidman-Duhovny, D.; Hammel, M.; Tainer, J. A.; Sali, A. FoXS , FoXSDock and MultiFoXS: Single-State and Multi-State Structural Modeling of Proteins and Their Complexes Based on SAXS Profiles. *Nucleic Acids Res.* **2016**, *44*, 424–429.
(62) Zhang, G. Y.; Liu, Y.; Ruoho, A. E.; Hurley, J. H. Structure of the Adenylyl Cyclase Catalytic Core (Vol 386, Pg 247, 1997). *Nature* **1997**, *388* (6638), 204.
(63) Donald, J. E.; Kulp, D. W.; DeGrado, W. F. Salt Bridges: Geometrically Specific, Designable Interactions. *Proteins Struct. Funct. Bioinforma.* **2011**, *79* (3), 898–915.
(64) Anantharaman, V.; Balaji, S.; Aravind, L. The Signaling Helix: A Common Functional Theme in Diverse Signaling Proteins. *Biol. Direct* **2006**, *1*, 25.
(65) Baskaran, P.; Heckler, E. J.; van den Akker, F.; Beuve, A. Identification of Residues in the Heme Domain of Soluble Guanylyl Cyclase That Are Important for Basal and Stimulated Catalytic Activity. *PLoS One* **2011**, *6* (11), e26976.

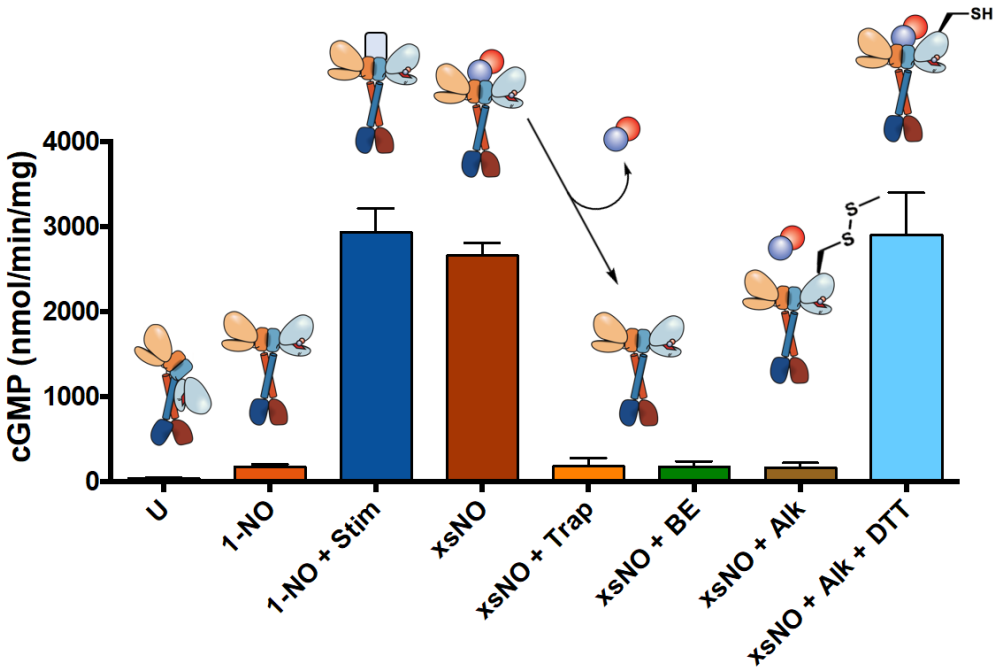
(66) Kang, Y.; Liu, R.; Wu, J.-X.; Chen, L. Structural Insights into the Mechanism of Human Soluble Guanylate Cyclase. *Nature* **2019**.
(67) Koglin, M.; Stasch, J. P.; Behrends, S. BAY 41-2272 Activates Two Isoforms of Nitric Oxide-Sensitive Guanylyl Cyclase. *Biochem. Biophys. Res. Commun.* **2002**, *292* (4), 1057–1062.
(68) Follmann, M.; Griebenow, N.; Hahn, M. G.; Hartung, I.; Mais, F. J.; Mittendorf, J.; Sch??fer, M.; Schirok, H.; Stasch, J. P.; Stoll, F.; et al. The Chemistry and Biology of Soluble Guanylate Cyclase Stimulators and Activators. *Angew. Chemie - Int. Ed.* **2013**, *52* (36), 9442–9462.
(69) Derbyshire, E. R.; Gunn, A.; Ibrahim, M.; Spiro, T. G.; Britt, R. D.; Marletta, M. A. Characterization of Two Different Five-Coordinate Soluble Guanylate Cyclase. *Biochemistry* **2008**, *47* (12), 3892–3899.
(70) Wales, J. A.; Chen, C.; Breci, L.; Weichsel, A.; Bernier, S. G.; Sheppeck II, J. E.; Solinga, R.; Nakai, T.; Renhowe, P. A.; Jung, J.; et al. Discovery of Stimulator Binding to a Conserved Pocket in the Heme Domain of Soluble Guanylyl Cyclase. *J. Biol. Chem.* **2018**, *293* (5), 1850–1864.
(71) Tesmer, J. J.; Sunahara, R. K.; Gilman, a G.; Sprang, S. R. Crystal Structure of the Catalytic Domains of Adenylyl Cyclase in a Complex with Gsα•GTPγs. *Science* **1997**, *278* (5345), 1907–1916.
(72) Hurley, J. H. Structure, Mechanism, and Regulation of Mammalian Adenylyl Cyclase. *J. Biol. Chem.* **1999**, *274* (12), 7599–7602.
(73) Guo, Y.; Suess, D. L. M.; Herzik, M. A.; Iavarone, A. T.; Britt, R. D.; Marletta, M. A. Regulation of Nitric Oxide Signaling by Formation of a Distal Receptor–Ligand Complex. *Nat. Chem. Biol.* **2017**, *13* (12), 1216–1221.
(74) Horst, B. G.; Marletta, M. A. Physiological Activation and Deactivation of Soluble Guanylate Cyclase. *Nitric Oxide - Biol. Chem.* **2018**, *77*.

CHAPTER 3

USING BIOCONJUGATIONAL CHEMISTRY TO CHARACTERIZE THE THREE-STEP ACTIVATION PATHWAY OF SOLUBLE GUANYLATE CYCLASE

Summary

The mammalian nitric oxide (NO) receptor, soluble guanylate cyclase (sGC) displays a unique activation profile in response to NO . Sub-stoichiometric amounts of NO relative to sGC result in a linear activation profile up to approximately 15% of maximal activity that is reached at one equivalent NO relative to sGC. The addition of excess NO increases the catalytic activity to 100% and removal of the excess NO returns the activity to 15% of maximal observed at one equivalent of NO. This biochemical behavior indicates that there are two distinct binding modes for NO on sGC: one to the heme cofactor and a second hypothesized to involve a specific cysteine residue. This chapter explores in detail the effect of addition of small bioconjugational agents on sGC activation. Methane methylthiosulfonate (MMTS) and N-ethyl maleimide (NEM), two cysteine reactive molecules, prevent excess NO from activating sGC from the 1-NO state to the maximally active state, while different lysine-labeling chemicals do not. However, the differential reactivity of these and other cysteine-reactive reagents influence the final activation profile of sGC. Structural work from Chapter 2 will be incorporated into the data here and future work will be proposed to more conclusively understand the effect of these bioconjugational agents on the activation of sGC and to further delineate the three-step activation process.



Research from parts of this chapter were performed by Carl C. Ward, Miriam Hood, Kimberly A. Houghton, and are thus indicated. This work is currently unpublished.

Introduction

Nitric oxide (NO) is a critical primary signaling molecule in cardiovascular and neuronal signaling.[1] The mammalian NO receptor, soluble guanylate cyclase (sGC), activates upon binding of NO to cyclize 5′-guanosine triphosphate (GTP) to 3′,5′-cyclic guanosine monophosphate (cGMP).[2–4] sGC can transduce the NO signal into cGMP when NO is in the picomolar to nanomolar concentration range, however the specific details of the activation of sGC by NO involve unique characteristics.[5]

sGC first binds NO through a heme B cofactor in the β H-NOX domain with a $K_d < 1.2 \times 10^{-12}$ M.[6] This interaction is specific for NO and essentially irreversible in a physiological setting; while carbon monoxide (CO) does bind relatively weakly to sGC, cellular CO concentrations are not high enough to maintain this interaction.[7] Biochemical quantification of sGC with only one equivalent of NO bound to sGC yield an enzyme that has only reached 15% of its maximal activity.[8,9] Full activity can be reached with the addition of excess NO and removal of this excess NO, either through a buffer exchange or a NO-specific trap returns the enzyme back to the 15% activity state.[9] NO remains bound to the heme after the removal of the excess NO. Previous work has studied the effect of different ligands bound to the heme of sGC in the absence and presence of excess NO. Derbyshire et al found that even with a non-native ligand bound to the heme (butyl isocyanate), excess NO still increases the activity of sGC.[10] These observations, both *in vitro* and in cells, have led to the hypothesis of a second NO binding site and the development of the three-state model for sGC activation: newly-translated sGC without any NO has low activity, sGC with NO only bound to the heme is only 15% active, while sGC with high nanomolar concentrations of NO has 100% activity.[11,12]

Discussion in Chapter 1 outlined previous work in the lab (Fernhoff et al.) that presented a hypothesis for the second site of NO on sGC and forms the basis for the experiments performed in this chapter.[13] The authors found that they could block the activation effect of excess NO on sGC by the treatment of sGC with S-methyl methanethiosulfonate (MMTS). Treatment of a MMTS-sGC sample with excess thiol (DTT) partially returned the activation by excess NO. This effect is independent of the presence or absence of oxygen, indicating that there was no redox process occurring. Fernhoff and colleagues concluded that the NO second site is a cysteine(s) on sGC, and that a direct attack of the thiolate onto NO to form a transient but stabilized radical anion was the basis for the second site interaction between NO and sGC. However, Chapter 2 demonstrated that sGC undergoes a large conformational change upon activation by excess NO. If MMTS modification of a cysteine residue is somehow preventing that conformational change from occurring, the same activation profile could exist. These hypotheses are not mutually exclusive; it is possible that MMTS is capping the second site cysteine at the same time it is preventing a large-scale conformational change.

Work in this chapter broadens the range of techniques as well as expands the experimental conditions used to probe this behavior of sGC. Homologues of sGC from lower organisms have been shown to possess the same three-step activation profile that mammalian sGC has.[14] Potential second site cysteines were narrowed by aligning *Manduca sexta* and *Rattus norvegicus* sGC and focusing only on conserved cysteines. Multiple types of protein mass spectrometry were brought to bear on this phenomenon, including a modification of activity-based protein profiling. Finally,

a screen of different bioconjugation agents solidified the role of cysteines as being important for activation for sGC. The work presented here, in combination with the work of Chapter 2, leads into substantial future directions that will continue to unravel the activation mechanisms of sGC.

Materials and Methods

Materials

All chemicals were purchased from commercial vendors and used without further purification, unless otherwise noted. Primers were obtained from Integrated DNA Technologies (Coralville, IA). PrimeSTAR Max DNA polymerase and TALON superflow resin was purchased from Takara Bio USA (Mountain View, CA). DNA purification kits were purchased from Qiagen (Germantown, MD). ExCell-420 media, sodium dithionite ($Na_2S_2O_4$), DNAse I from bovine pancreas, β-mercaptoethanol, sodium phosphate, dibasic (Na_2HPO_4), sodium carbonate (Na_2CO_3), potassium phosphate (K_2HPO_4), guanosine 5′-triphosphate sodium salt (GTP) (>95%, HPLC), zinc acetate ($Zn(CH_3CO_2)_2$), methyl methanethiosulfonate (MMTS, 97%), Iodoacetic Acid (IAC), Iodoacetamide (IAM), and N-ethyl maleimide (NEM) were purchased from Millipore Sigma (Burlington, MA). 4-(2-aminoethyl)benzenesulfonyl fluoride hydrochloride (AEBSF) was purchased from Research Products International (Mount Prospect, IL). Isopropyl β-D-1-thiogalactopyranoside (IPTG), 4-(2-hydroxyethyl)piperazine-1-ethanesulfonic acid (HEPES), Slida-A-Lyzer mini dialysis devices (10,000 MWCO), and Sulfo N-hydroxysuccinimide acetate (s-NHS-Ac) were purchased from Thermo Fisher Scientific (Waltham, MA). Benzamidine, sodium chloride (NaCl), magnesium chloride ($MgCl_2$), and glycerol were purchased from VWR Life Science (Visalia, CA). Dithiothreitol (DTT) was purchased from Bachem (Bubendorf, Switzerland). Imidazole was purchased from Oakwood Chemical (West Columbia, SC). Vivaspin spin concentrators were purchased from Sartorius (Concord, CA). BioSpin 6 desalting columns were purchased from BioRad (Hercules, CA). Diethylamine NONOate (DEA NONOate), PROLI NONOate, S-nitrosoglutathione (GSNO), and YC-1 were purchased from Cayman Chemical Company (Ann Arbor, MI). Carbon monoxide (CO, >99%) gas was purchased from Praxair Inc. (Danbury, CT). Isatoic anhydride (IsaAnh) was a generous gift from Matthew Francis's lab.

Construction of Plasmids

Rattus norvegicus (*Rn*) sGC α1 (Uniprot: P19686) and β1(Uniprot: P20595) genes were cloned into pFastBac to generate pFasBac_Rn_α1_His_6 and pFastBac_Rn_β1 previously.[15] *Manduca sexta* (*Ms*) sGC α1 (Uniprot: O77105) and β1(Uniprot: O77106) genes were cloned into pFastBac to generate pFastBac_*Ms*_sGC_α1_His6 and pFastBac_*Ms*_sGC_β1 (Chapter 2). For mutants of *Rn* sGC, two- or three-piece Gibson Assembly fragments were amplified from pFasBac_Rn_α1_His_6 or pFastBac_Rn_β1, and then ligated together with Gibson Assembly. The genes were transposed into a baculovirus bacmid using DH10Bac-GFP cells. The bacmid was isolated, validated for gene transposition by amplification and sequencing (UC Berkeley DNA Sequencing Facility), and then transfected into SF9 cells (Gibco) to generate recombinant baculovirus for both genes.

Protein Expression and Purification

The following steps were used to express and purify both *Rn* and *Ms* sGC. SF9 cells were maintained in monolayer and in suspension in ExCell-420 media at 27 °C (cells were shaken at 135 rpm). The recombinant baculovirus was amplified until the titer was greater than 1×10^8

cfu/mL. Five liters of SF9 cells were coinfected with 50 mL of amplified virus per liter and allowed to express for 72 hours. Cells were spun down at 4300 *g* for 20 minutes, snap frozen in liquid nitrogen, and stored at –80 °C.

Two different purification procedures were followed, one for enriched lysate assays and one for sGC purified to homogeneity:

The following steps to enrich sGC lysate were performed at 4 °C. Cell pellets from a 50 mL SF9 cell expression were thawed in ice water and resuspended in 1 mL of lysis buffer: Buffer A (50 mM Na_2HPO_4, pH 8.0, 200 mM NaCl, 1 mM imidazole, 1 mM benzamidine, 5% (v/v) glycerol, 0.22 μm filtered) supplemented with 10 mM benzamidine, 1 mM AEBSF, 0.5 mg/mL bovine DNAse I, and 5 mM β-mercaptoethanol. Cells were lysed by sonication with a microtip adapter at power level 1.5 for the following program: 15 seconds on followed by 15 seconds off for a total process time of 30 seconds. Cell debris was spun down in a 2 mL microcentrifuge tube for 1 hr, 21,130 *g*. To the clarified lysate, 50 μL of TALON superflow beads pre-equilibrated with Buffer A supplemented with 5 mM β-mercaptoethanol were added. The beads and clarified lysate mixture was allowed to rock for 30 min. The beads were pelleted with a gentle spin in the microcentrifuge, 1 minute, 1,000 *g*, and the supernatant was saved. The beads were washed twice with 500 μL of Buffer A supplemented with 5 mM β-mercaptoethanol. sGC was eluted off the beads with 100 μL Buffer B (50 mM Na_2HPO_4, pH 8.0, 200 mM NaCl, 150 mM imidazole, 1 mM benzamidine, 5% (v/v) glycerol, 0.22 μm filtered) supplemented with 5 mM β- mercaptoethanol. Enriched lysate was snap frozen in liquid nitrogen and stored at –80 °C.

The following steps to purify sGC to homogeneity were performed at 4 °C. Cell pellets from a 5 L SF9 cell expression were thawed in ice water and resuspended in lysis buffer. Cells were lysed in a bead beater (BioSpec) with 0.5 mm glass beads. The resulting cell lysate was clarified by spinning at 4,300 *g* for 5 minutes, followed by spinning at 158,000 *g* (42,000 RPM) for 2 hours (Ti45 rotor, Beckman Coulter). The lysate was passed through a column containing 2 mL TALON superflow at 0.5 mL/min, and the flow-through was collected. The column was washed with 15 column volumes of Buffer A supplemented with 5 mM β-mercaptoethanol at 0.5 mL/min. The protein was eluted with 10 CV of Buffer B. Fractions with yellow color were concentrated to < 2 mL using a 30,000 MWCO cutoff spin concentrator, supplemented with 5 mM DTT and 5 mM EDTA, and stored overnight. The next day, the sample was diluted to 9 mL with Buffer C (25 mM triethanolamine, 25 mM NaCl, 5 mM DTT, 5% (v/v) glycerol, 0.22 μm filtered), and applied to a POROS HQ2 anion exchange column (Thermo Fisher Scientific). The column was washed with 3 CV of Buffer C, and then a gradient to 50% Buffer D (25 mM triethanolamine, 750 mM NaCl, 5 mM DTT, 5% (v/v) glycerol, 0.22 μm filtered) was established over 17 CV at 0.5 mL/min. Fractions with purified sGC were concentrated to 5–50 μM and stored in liquid nitrogen. A typical yield for this expression and purification procedure is 100-150 μg sGC per liter of insect cells.

Sample Preparation and Labeling for Mass Spectrometry and Activity Assays

Sample preparation was similar for both mass spectrometry and activity assays (differences noted in later sections). Samples were reduced in an anaerobic glove bag with 5 mM $Na_2S_2O_4$ for 15 minutes, and then desalted with a BioSpin6 column equilibrated with either Buffer E (50 mM HEPES, pH 7.5, 150 mM NaCl, 5% (v/v) glycerol, 0.22 μm filtered) for labeling, activity assays, and intact protein mass spectrometry, or Buffer F (50 mM Na_2HPO_4, pH 8.0, 200 mM NaCl, 5%

(v/v) glycerol, 0.22 μm filtered) for ABPP experiments. For samples that were to be labeled with cysteine alkylating agents, excess DTT was removed from the sample in a variety of ways: five dilutions with Buffer E and concentrations with a Vivaspin 500 (30,000 MWCO) spin filter, or two 30-minute rounds of dialysis of the reduced sGC against Buffer E with a sample to dialysate ratio of 1:10,00. It has been found that the dialysis method retained the highest amount of sGC. sGC with 1-NO and xsNO were prepared by first adding PROLI NONOate to 50 μM, based on 2 moles of NO released per mole of PROLI NONOate. This sample was then buffer exchanged into Buffer E through a BioSpin6 column to generate the 1-NO state. To generate the xsNO state, PROLI NONOate was added back to a portion of the 1-NO sample and allowed to equilibrate for 5 minutes. The protein concentration was determined using the reduced heme Soret peak at 433 nm (148,000 M^{-1} cm^{-1}).[14] YC-1 was added from a 100x stock solution in DMSO to a final concentration of 150 μM (final DMSO concentration 1%). When called for, MMTS, IAC, IAM, s-NHS-Ac, and IsaAnh were diluted from 100x stocks into reduced sGC samples for final concentrations between 10 and 200 μM, depending on the specific experiment. Samples were allowed to label for 30 minutes to 1 hour.

Activity Assays and Quantification

Maximum activity for *Rn* and *Ms* sGC were measured by quantifying the amount of cGMP produced in duplicate endpoint activity assays. Activity assays were conducted at 25 °C and pH 7.5 in Buffer E, supplemented with 5 mM DTT and 5 mM $MgCl_2$ for experiments without labeling agents and only 5 mM $MgCl_2$ for experiments with labeling agents. Reactions were initiated with 1.5 mM GTP and timepoints were removed from the reaction mixture and quenched with 125 mM $Zn(CH_3CO_2)_2$, followed by 125 mM $Na_2(CO_3)$ to adjust the pH to 7. Samples were frozen at –80 °C until quantification. Quenched assays were thawed, and the zinc precipitate was spun down for 10 min at 23,150 g. The reactions were diluted by one to three orders of magnitude, and the cGMP was quantified in duplicate using an extracellular cGMP Enzyme Linked Immunosorbent Assay, following the manufacturer's instructions (Enzo Life Sciences). Concentrations of cGMP were determined from a standard curve, generated over 0.16–500 pmol/mL. Initial rates were calculated from the linear phase of the time course, where 5–10% of the GTP substrate was consumed.

Intact Protein Mass Spectrometry

Purified proteins were treated as outlined in the above Sample Preparation section. Prior to intact protein mass spectrometry, samples were buffer-exchanged into 25 mM HEPES, pH 7.5, 25 mM NaCl using three rounds of dilution and concentrations in a Vivaspin 500 (30,000 MWCO) spin concentrator. Samples were filtered with a 0.22 μM spin filter (Millipore). The final sample concentration was approximately 1 μM. Samples of 2 μL were injected onto and separated over a ProSwift column (ThermoFisher Scientific) with an Agilent 1200 series high pressure liquid chromatography (HPLC) system, and subsequently analyzed by an Agilent 6224 time-of-flight (TOF) mass spectrometer with a Turbospray ion source in positive ion mode.

Sample Preparation for Activity Based Protein Profiling Workflow

Purified proteins were treated as outlined in the above Sample Preparation section, and diluted into Buffer F. Protein was concentrated to a final concentration of 2 μM and then samples were transferred to C. Ward for further preparation similar to previously published protocols.[16] Briefly, samples were incubated with 100 μM of alkylating agent at room temperature for one hour. Samples were then acetone precipitated, resuspended in methanol, and combined according to

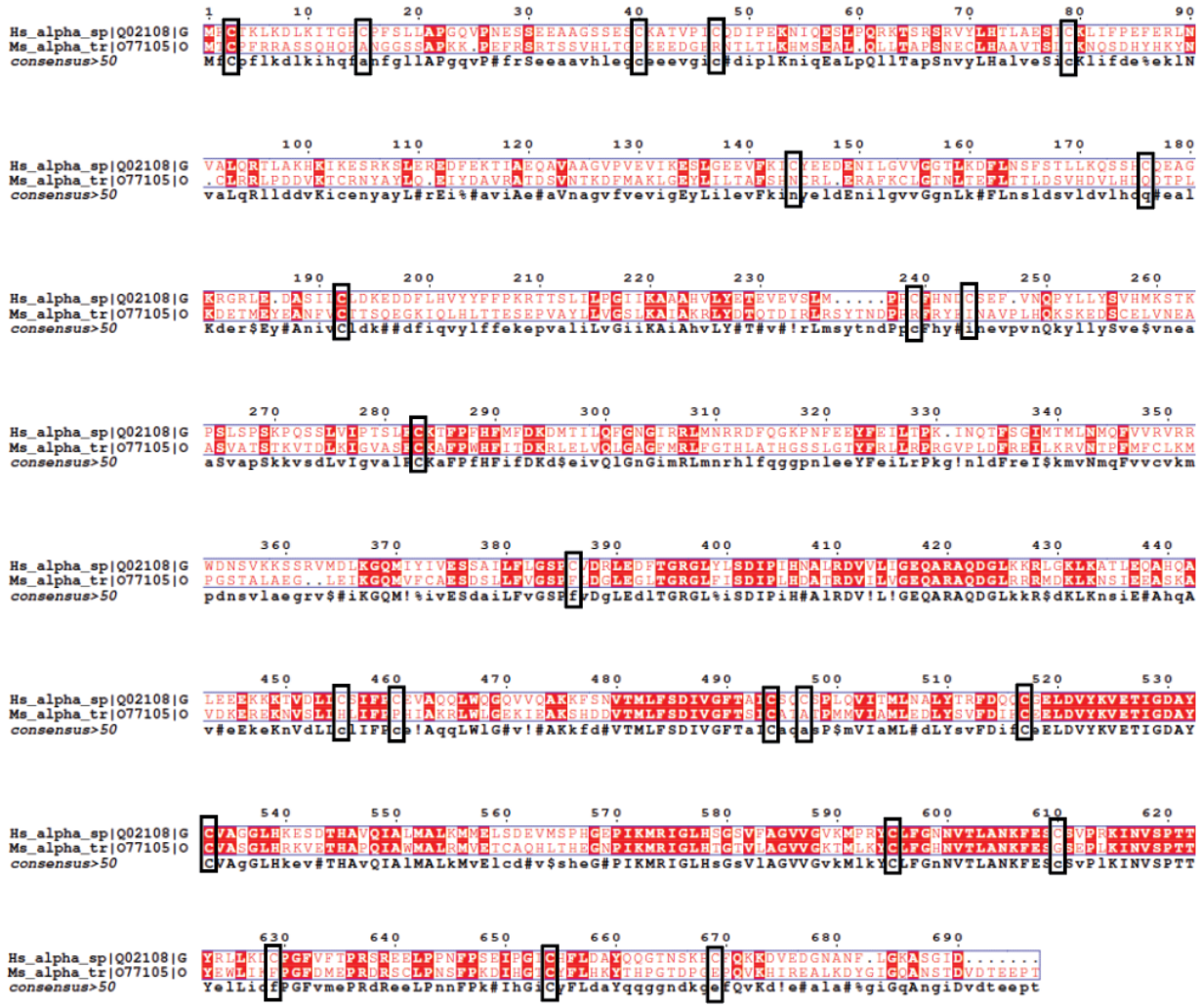
experimental design. Samples were digested with trypsin overnight at 4 °C and further prepared for mass spectral analysis.

Results and Discussion

Given the hypothesis of cysteine involvement in second NO binding site of sGC and the challenge of screening large numbers of sGC variants, sequence alignments and mass spectrometry were employed in attempts to narrow down the potential critical cysteine residue regulating the activity between the 1-NO and excess NO state. Attempts made to express *Ms* sGC in *E. coli* as has been reported previously were unsuccessful, and thus these alternate approaches were explored.[14]

Narrowing the Cysteine List via Sequence Alignments

In the first approach, sequence alignments were used to examine the number of conserved cysteines between three sGC homologues known to possess a three-state activation profile. Results presented in Chapter 2 illustrated that *Ms* sGC (in addition to *Hs* and *Rn* sGC) possess three different states of activity (U, 1-NO, and xsNO) and retains the ability for YC-1 to activate the 1-



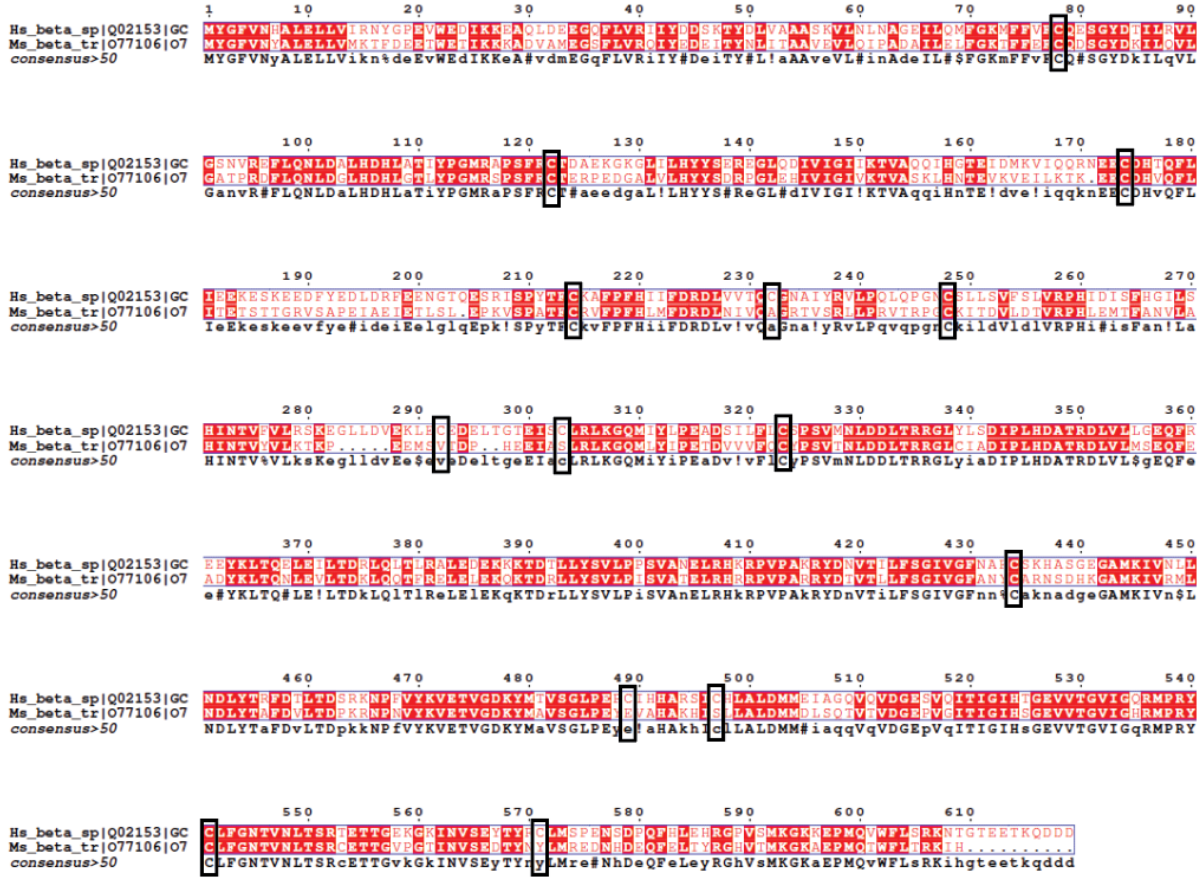


Figure 3.1 – Multiple sequence alignments of *Hs* and *Ms* sGC. The α1 and β1 genes are aligned separately. All cysteines of *Hs* sGC are boxed by black rectangles. Only 16 of these residues are conserved between *Hs* and *Ms* sGC sequences.

NO state to maximal activity. As *Hs* and *Rn* sGC 89.4% and 98.9% identical for the α1 and β1 sequences respectively, only the *Hs* sGC sequence will be used for comparison with *Ms* sGC (the same cysteines are conserved when *Rn* sGC is used for alignments). Figure 3.1 depicts the multiple sequence alignment between the sequences of *Hs* and *Ms* sGC α1 and β1 sequences. Black boxes indicate sites where *Hs* sGC contains a cysteine residue. Of the 37 cysteine residues present in the *Hs* sGC sequence, only 16 of those residues are conserved between *Hs* and *Ms* sGC, indicating the importance cysteine residues at those positions.

It was further reasoned that a second site cysteine would be involved either stabilizing or interrupting a conformational change of sGC between the U, 1-NO, and excess NO states. Thus, this list of 16 cysteine residues was further narrowed to include cysteines whose peptides showed increases or decreases in exchange rates based on Hydrogen Deuterium Exchange Mass Spectrometry (HDX-MS).[17] Figure 3.2A shows the *Ms* sGC heterodimer showing those 16 conserved cysteines in purple with ribbons colored according to the differences in deuterium exchange rates as measured by HDX-MS between the excess NO state and the unliganded state, with larger exchange rates in the excess NO state colored more red. Cysteines αC532 and βC248

were selected based on their location adjacent to or within peptides on sGC that had decreased exchange when comparing the unliganded and excess NO state. These residues were mutated conservatively to serine, expressed in SF9 cells, and purified to homogeneity.

By mutating the proposed second site cysteine to a serine residue, the interaction between NO and the critical cysteine would not be able to form and stabilize the active state, thus rendering the protein insensitive to excess NO. Additionally, treatment of sGC with MMTS should not have any effect on activity because the key cysteine has been removed. Figure 3.2B shows the activation profile for WT, αC532, and βC248 variants of *Rn* sGC under eight different conditions. All three proteins retain the ability to be activated by xsNO and stimulated by YC-1 in the 1-NO state. Additionally, none of the variants were stimulated under any conditions with MMTS present. These assays indicate that neither αC532 nor βC248 are solely responsible for the second NO binding site and other cysteines present in sGC are being modified by MMTS to prevent the stimulation by excess NO.

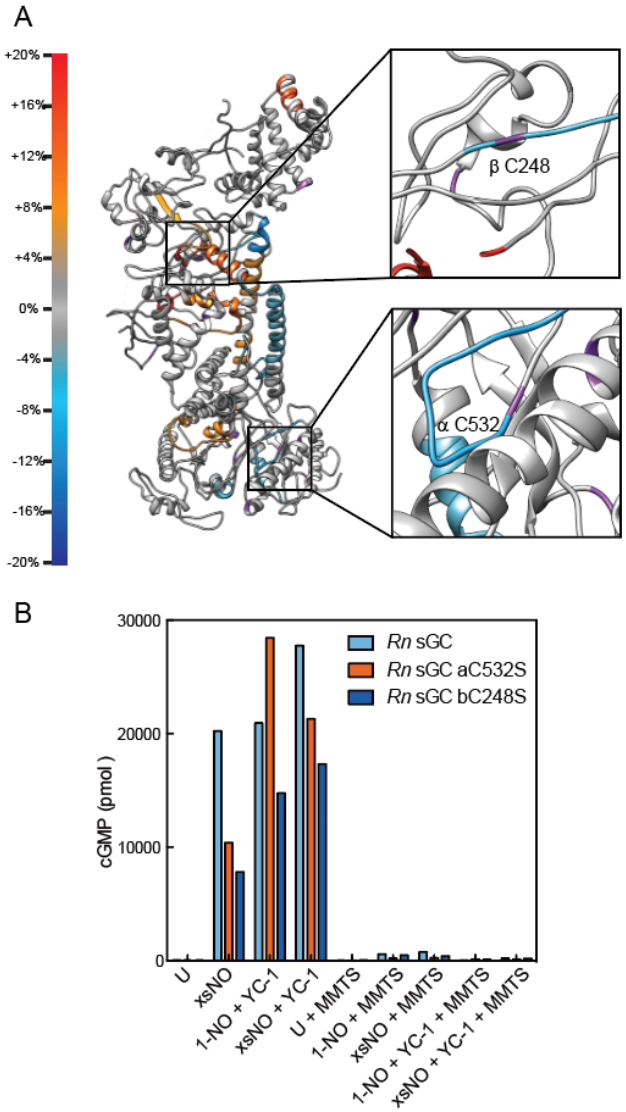


Figure 3.2 – HDX-MS targeted selection of conserved cysteines A) HDX-MS data overlaid on the inactive structure of *Ms* sGC (HDX-MS from ref 17, *Ms* sGC structure from Chapter 2). Cysteines are colored purple. B) End point activity assays of *Rn* sGC and indicated variants with different small molecule stimulators and inhibitors. Contrary to previously published data, YC-1 did not rescue *Rn* sGC.[13]

To screen a larger set of these cysteine residues that were conserved between *Hs* and *Ms* sGC, 14 of the remaining 16 cysteine residues were mutated to serines and the lysate of SF9 cells was enriched to screen for inhibition of activation. The two cysteines that were not screened were αC3 (because the N-terminus of alpha is known to be processed in some homologues and is thus not conserved in all mature proteins) and β541C (as this cysteine helps guanylate cyclases discriminate between GTP and ATP as substrates). As a result of performing the assays in enriched lysate, the 1-NO state could not be interrogated in this set of experiments. Instead, excess NO with MMTS, excess NO, and excess NO with YC-1 were examined. Wildtype sGC is predicted to have low activity with excess NO and MMTS, high activity with excess NO, and high activity with excess NO with YC-1. The activation profile of a variant with a second site cysteine mutated to a serine should only vary in the excess NO condition, where it would now be predicted to be low activity. As seen in Figure 3.3A, the enriched lysate assay has a much lower dynamic range than protein purified to homogeneity, however the wildtype enzyme does show activation by excess NO and excess NO with YC-1 over excess NO with MMTS. A few variants screened show decreased activation with excess NO, namely αC653S and βC78S. The βC78S had been purified previously and shows decreased heme incorporation.[18] Thus, the αC653S variant was expressed and purified to homogeneity. Figure 3.3B displays the activation profile of WT and αC653S *Rn* sGC, which both clearly show sensitivity to excess NO. That excess NO activates αC653S *Rn* sGC shows that, similar to αC532 and βC248 selected form HDXMS analysis, αC653 by itself does not constitute the second NO binding site.

Increasing the number of sequences under examination can illuminate trends overlooked by comparing only two sequences. A BLAST search of the *Rn* β H-NOX domain (residues 1-191) returns 2,186 sGC sequences that are represented in Figure 3.4A as a sequence similarity network (made with the Enzyme Function Initiative's EFI-EST tool).[19] SSNs can be thought of as visual representations of a phylogenetic trees. The user selects an alignment score that is equivalent to a BLAST expectation value to set a threshold value for how similar two sequences (nodes) must be for an edge to be drawn between them. The alignment score can be manually increased until proteins with similar functions are contained within the same clusters. Here, an alignment score of 130 efficiently separates mammalian and insect α subunits and β subunits from O_2-sensitive sGCs and other sGCs from algae and choanoflagellates. Extraction of the sequences within the clusters that contain *Rn*, *Hs*, and *Ms* sGC α1 (orange) and β1 (blue) sequence yield a larger set of proteins that are predicted to share the same function. Sixty α1 and 215 β1 sequences were extracted and aligned separately with Clustal Omega. A portion of the β1 alignments are displayed in Figure 3.4B as a WebLogo, in which the height of the amino acid letters indicates the fraction of sequences in the alignment that have that residue in that position.[20] Cysteines boxed in red are conserved in *Hs*, *Rn*, and *Ms*, however they are not strictly conserved with greater than 5% of the expanded sequence alignment. In total, five cysteines are not strictly conserved in this analysis, bringing the final count of strictly conserved cysteines down to 11: αC191, αC282, αC493, αC516, αC532, αC594, βC78, βC122, βC323, βC434, and βC541 (active site GTP-selective residue). If the second NO binding site is a cysteine residue, it is likely to be contained within this set of cysteines.

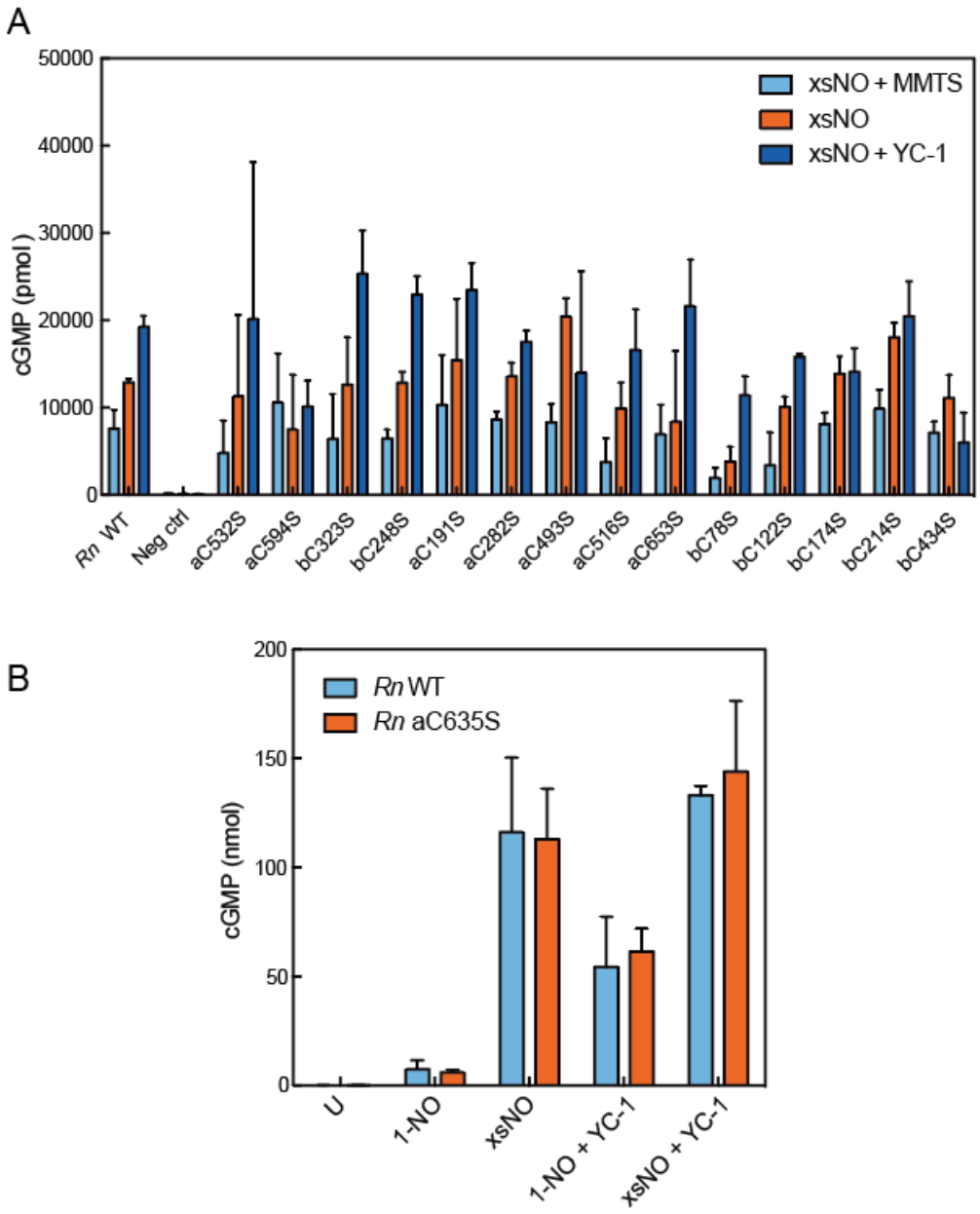


Figure 3.3 – Full screen of targeted cysteines using a lysate assay. A) An enriched lysate screen of *Rn* sGC WT and selected variants. Each variant was treated with excess NO and then split into three fractions: MMTS to simulate 1-NO activity, nothing further to simulate excess NO, and YC-1 to simulate full activity. B) The most promising mutant that was not known to interrupt protein function was selected to characterize in more detail. *Rn* αC635S has a similar activation profile to *Rn* sGC WT.

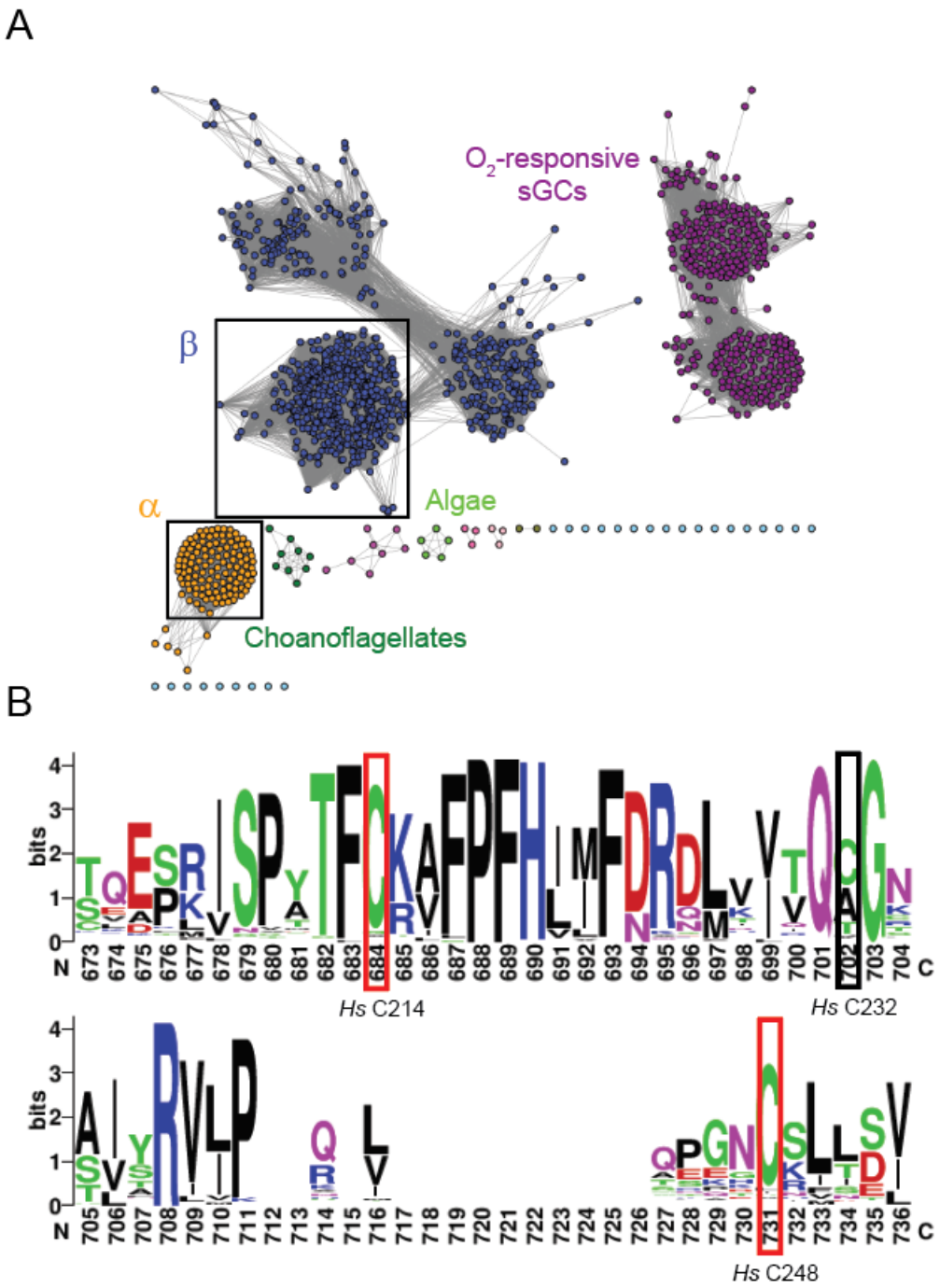


Figure 3.4 – Expansion of sequences considered using bioinformatics. A) Sequence Similarity Network generated from a BLAST of the *Rn* sGC β H-NOX domain with an alignment score of 130. Sequences were only included in the SSN if they were between 500 and 900 amino acids in length. The α-like sequences are colored in orange and the β-like sequences are colored in blue (boxed). O_2 sensitive sGCs, choanoflagellate sGCs, and algae sGCs are indicated (see Chapter 4 and 5). The blue cluster that extend off of the boxed β cluster are atypical sGCs and were not included in the subsequent analysis. B) WebLogo generated based on the sequences extracted from the β-like sequences. Red boxes indicated the cysteine is conserved in both *Rn* and *Ms* sGC, black boxes indicated the cysteine is present only in *Hs* sGC. βC214 and βC248 are conserved between *Hs* sGC and *Ms* sGC, however within the 215 sequences extracted from the SSN these cysteines are not strictly conserved.

Narrowing the Cysteine List via Proteomics

The second approach that was used in attempts to narrow down the location of the second site cysteine was a modified Activity Based Protein Profiling (ABPP) approach. Briefly, ABPP relies on the differential reactivity of amino acids in order to extract peptides that contain hyper-reactive residues. For example, this approach has worked well to identify reactive cysteines in cysteine proteases within a proteome. This is because as the ratio of two different probes is varied, the hyper-reactive cysteines in the active site of cysteine proteases are insensitive to probe concentration, whereas non-reactive cysteines are probe concentration dependent. This work was done in collaboration with C. Ward from the Nomura Lab at UC Berkeley. Figure 3.5A shows a schematic of one potential design of the experiment as it relates to sGC. We hypothesized that the second site cysteine would have to be quite reactive to sense NO at such low concentrations and form the radical anion NO adduct. By varying the ratio of deuterated to non-deuterated NEM, the most reactive cysteines could be extracted. Figure 3.5B depicts the area ratios from different peptides. We would expect the most reactive cysteines to be clustered together with a low log ratio, however the data was not clear in this regard. Unexpectantly, cysteine reactivity was variable across different runs. Because of this and technical difficulties surrounding amount of protein to generate a good signal, this experimental approach was not pursued further.

An alternate hypothesis that was also explored using the same sequencing mass spectrometry techniques is that once the second site cysteine interacts with NO it would not interact with alkylating agents, in other words using NO as a blocking agent. In this experimental set up, the prediction is that there would be one cysteine where the ratio of labeling in the U state would be much higher than labeling in the excess NO state. Two different probes, N-(5-hexyn-1-yl)acetamide (IAyne) and 1-(5-Hexyn-1-yl)-1H-pyrrole-2,5-dione (Maleimide-yne) were used to explore slightly different reactivities and steric effects. The data from this set of experiments is shown in Figure 3.5C and does not indicate a single cysteine that has a higher area ratio in the unliganded state than the excess NO state. In contrast, the data show that all cysteines sequenced are labeled more in the excess NO state. This result makes sense in light of the data presented in Chapter 2, where excess NO gives rise to a large conformational change of sGC to potentially expose more cysteines to alkylating agents. However, these data do not help to prioritize a specific set of cysteines that potentially comprise the second NO binding site.

Activity Assays with Bioconjugational Agents and YC-1

At this point, alternate hypotheses were explored for the observation of MMTS-induced inhibition of the excess NO state of sGC. It was hypothesized that if excess NO and YC-1 were indeed stimulating sGC through different mechanisms, then sGC treated with excess NO, MMTS, and YC-1 would show an increase in activity, as has been seen before.[13] Figure 3.6 shows the activation profile of both *Rn* and *Ms* sGC in the 1-NO and excess NO states with and without YC-1. In the conditions assayed here, YC-1 does not rescue sGC activity when MMTS is present in either the 1-NO or excess NO state.These data indicate that either the conditions where YC-1 rescues a MMTS-treated sample of sGC are very specific, MMTS is preventing a conformational change, or perhaps excess NO and YC-1 are acting through the same mechanism to activate sGC.

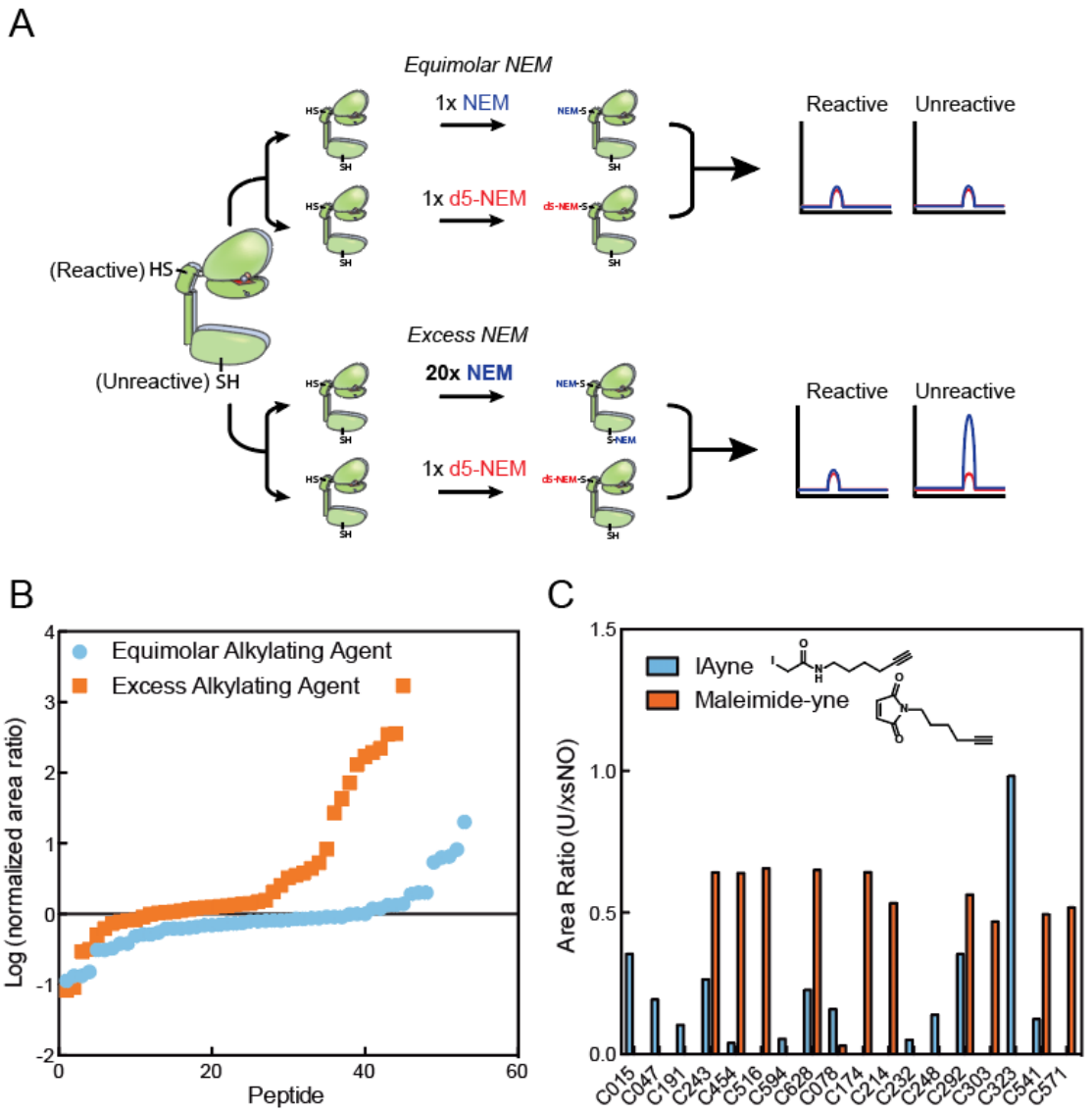


Figure 3.5 – Activity Based Protein Profiling approaches to narrow cysteine list. A) Overall scheme of Activity Based Protein Profiling approach to look for reactive cysteines in sGC. In this example two sGC cysteines are shown, one that is reactive and the other that is not. If equal molar ratios of heavy and light NEM are added to two samples of sGC in substoichiometric amounts compared to the number of sGC cysteines, only the reactive cysteines should be labeled, and they should be labeled in the equivalent ratios. If unreactive cysteines are labeled, they should also be labeled in equivalent ratios (top). However, if an excess of the light NEM is used, reactive cysteines should still be labeled equivalently while unreactive cysteines should have a larger area corresponding to reaction with light NEM (bottom). B) The graph shows the observed peptides ordered such that the peptides with the smallest area ratio on the left, indicating more reactive cysteines and larger area on the right. This trend does not hold if organized by cysteine number (e.g. the cysteines that are most reactive with equimolar alkylating agents do not appear to be most reactive in the case with the excess alkylating agents). C) The graph shows the results of the second strategy in which NO imagined to be used as a block for the second site cysteine. The y axis depicts the area of modified peptides in the mass spectrum chromatogram modified in the unliganded state divided by the area of modified peptides in the mass spectrum in the excess NO state. We would predict that the second site cysteine would have an area ratio of much greater than one, however in general we saw that more cysteines were reacting with both alkylating agents in the excess NO state. This is most likely attributed to the larger conformational change that occurs between the unliganded state compared to the excess NO state which potentially allows for more cysteines to be modified.

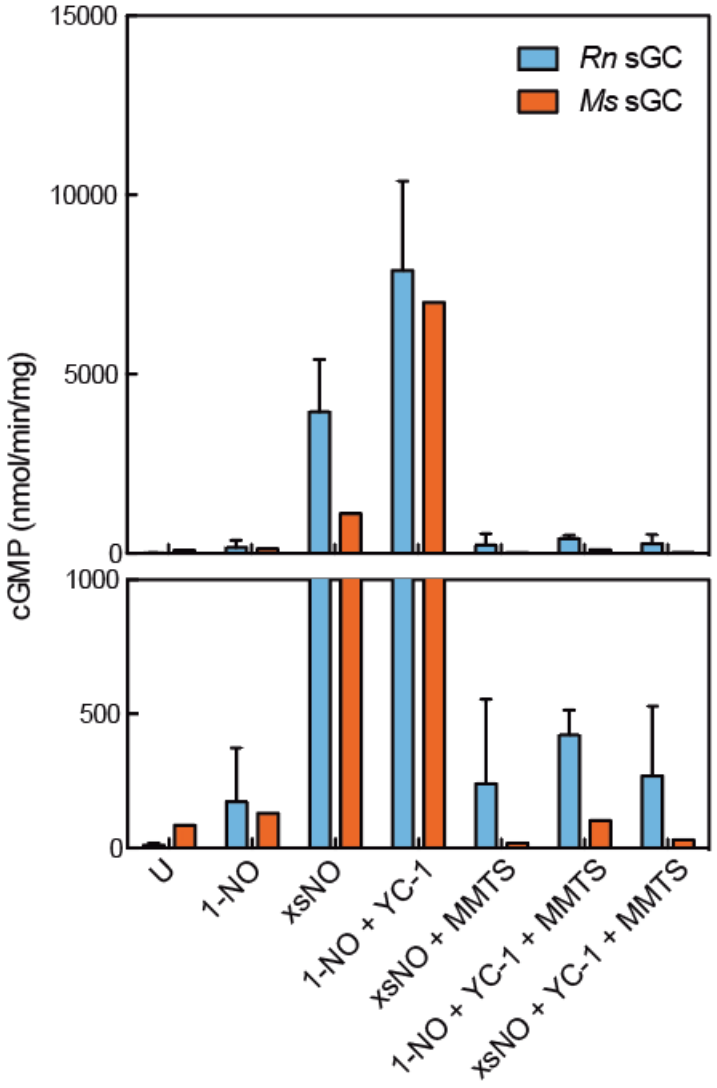


Figure 3.6 – sGC activity with MMTS and YC-1. Activity data for both *Rn* and *Ms* sGC depicting the effect of small molecule stimulators and inhibitors on activity. In contrast to previously published data, YC-1 does not stimulate MMTS-treated *Rn* or *Ms* sGC.[13] It is possible that under some conditions, YC-1 is able to stimulate a MMTS-treated sGC but it is not true for the conditions assayed here.

Perhaps MMTS was not modifying the second site cysteine, but in fact was preventing the conformational change that allowed sGC to activate. This hypothesis would be supported if modifying agents that targeted other amino acids also inhibited activation of sGC by excess NO. To explore the hypothesis that MMTS was inhibiting a conformational change, different bioconjugational agents that modify both cysteine and lysine residues were used to modify sGC. The bioconjugation reactions used in this section as well as the mass change for the modified amino acid are depicted in Figure 3.7A. Preliminary activity assays with MMTS, NEM, s-NHS-Ac, and IsaAnh indicated that the inhibition of excess NO-stimulated activity was indeed constricted to cysteine modifying agents (Figure 3.7B). This assay indicates that perhaps the modification of other residues does not prevent excess NO-stimulated activity, but it is still possible that a cysteine in a critical position that is modified by MMTS and NEM is preventing a

crucial conformational change. An alternate explanation for the differential activity is that different stoichiometries and reactivities of the modifying agents to amino acids results in a different average number of amino acids labeled per heterodimer of sGC.

To quantify on average how many modifications are found on sGC after treatment with different bioconjugational agents, an assay was performed to measure both intact protein mass spectra and endpoint activity in parallel. The final concentration of *Rn* sGC in these labeling reactions was 1 μM, so based on the number of cysteines and lysines in the proteins gives final concentrations of 34 μM cysteine and 78 μM lysine. Figure 3.8A displays a deconvoluted mass spectrum that shows the mass changes for sGC subjected to the indicated final concentration of labeling agent, with the largest observed change noted in number of molecules per subunit. Interestingly, the average mass change for sGC treated with MMTS was much larger than expected, perhaps indicating that the MMTS reagent was more concentrated than calculated. All other agents modified sGC in a range that was closer to the expected average number of molecules per subunit. Figure 3.8B shows the results of the activity assay performed in parallel. MMTS inhibited sGC in both excess NO and excess NO with YC-1, however neither IAM nor IAC showed this same inhibition despite both agents modifying sGC. Both IAM and IAC react slower than MMTS and NEM, perhaps indicating either not enough time was allowed for the labeling reactions to progress to completion. Thus, titrations of all bioconjugational agents must be performed to see if the inhibition of sGC activity occurs at the same number or different number of residues modified per subunit. A first titration of sGC activity with MMTS is shown in Figure 3.8C, where maximal inhibition occurs between 50 and 100 μM MMTS for 2 μM sGC.

To explore the possibility of using the differential reactivities of MMTS and IAM to our advantage, two rotations students in the Marletta lab M. Hood and K. Houghton attempted to label sGC first with IAM followed by MMTS and compared the activation of excess NO with a sample that was reacted with IAM followed by IAM and MMTS followed by MMTS. Figure 3.9 shows that it is possible to label sGC first with IAM followed more IAM and still retain NO-stimulated sGC activity, however addition of MMTS under any conditions prevents the NO-stimulated activity. This line of research is similar to what was attempted with the ABPP where we imagined NO functioning as a blocking agent, however IAM forms a covalent bond to the cysteine. These double addition type experiments hold great promise for the elucidation of critical cysteines in sGC.

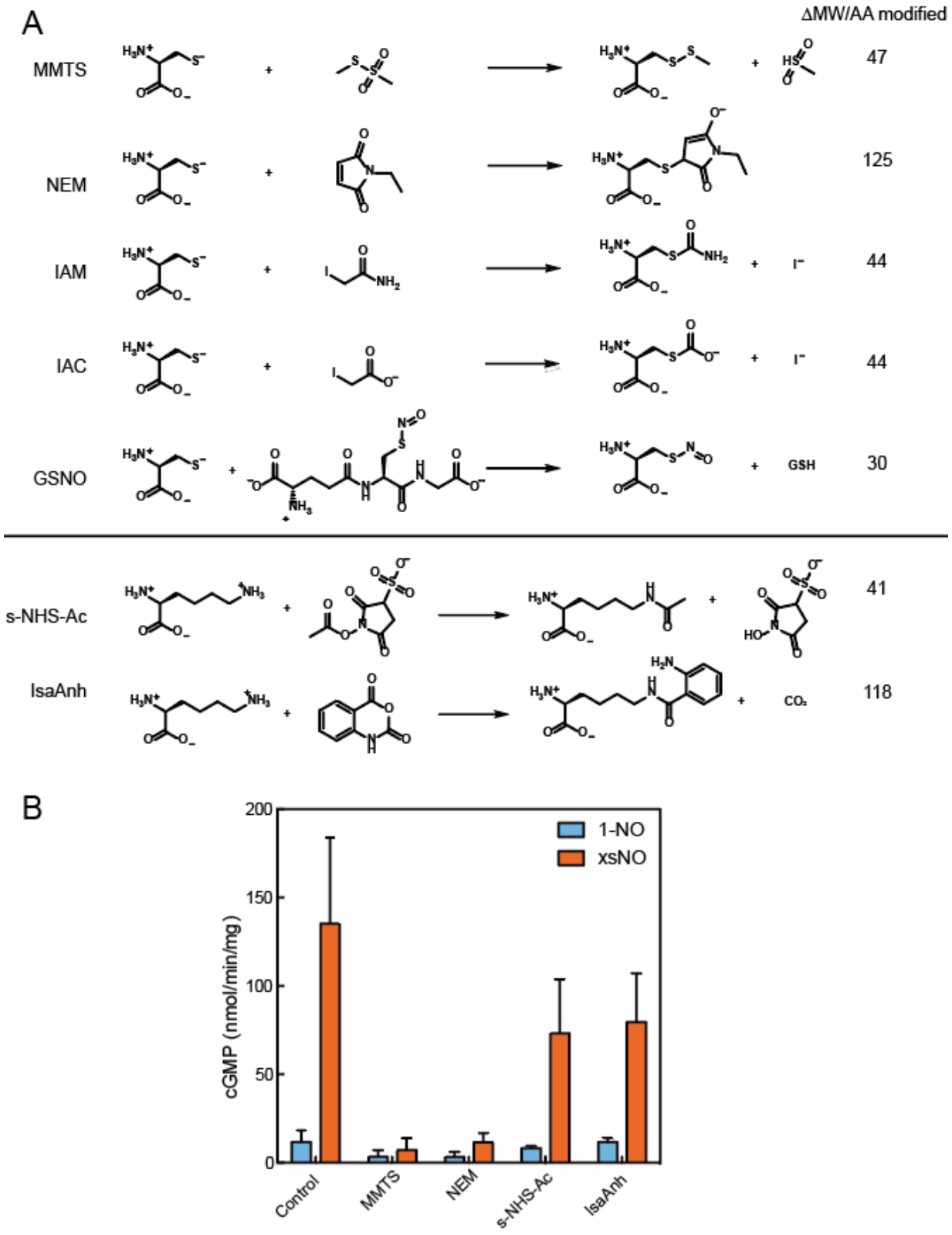


Figure 3.7 – Expansion of small molecule bioconjugational agents used on sGC. A) The reactions of small molecule bioconjugational agents used in this chapter is shown, along with the mass change of the amino acid they modify (right). B) Activity data of *Rn* sGC in the 1-NO and excess NO state. Both cysteine labeling agents abrogated excess NO-stimulated activity of *Rn* sGC, while labeling with lysine labeling agents did not have this same effect.

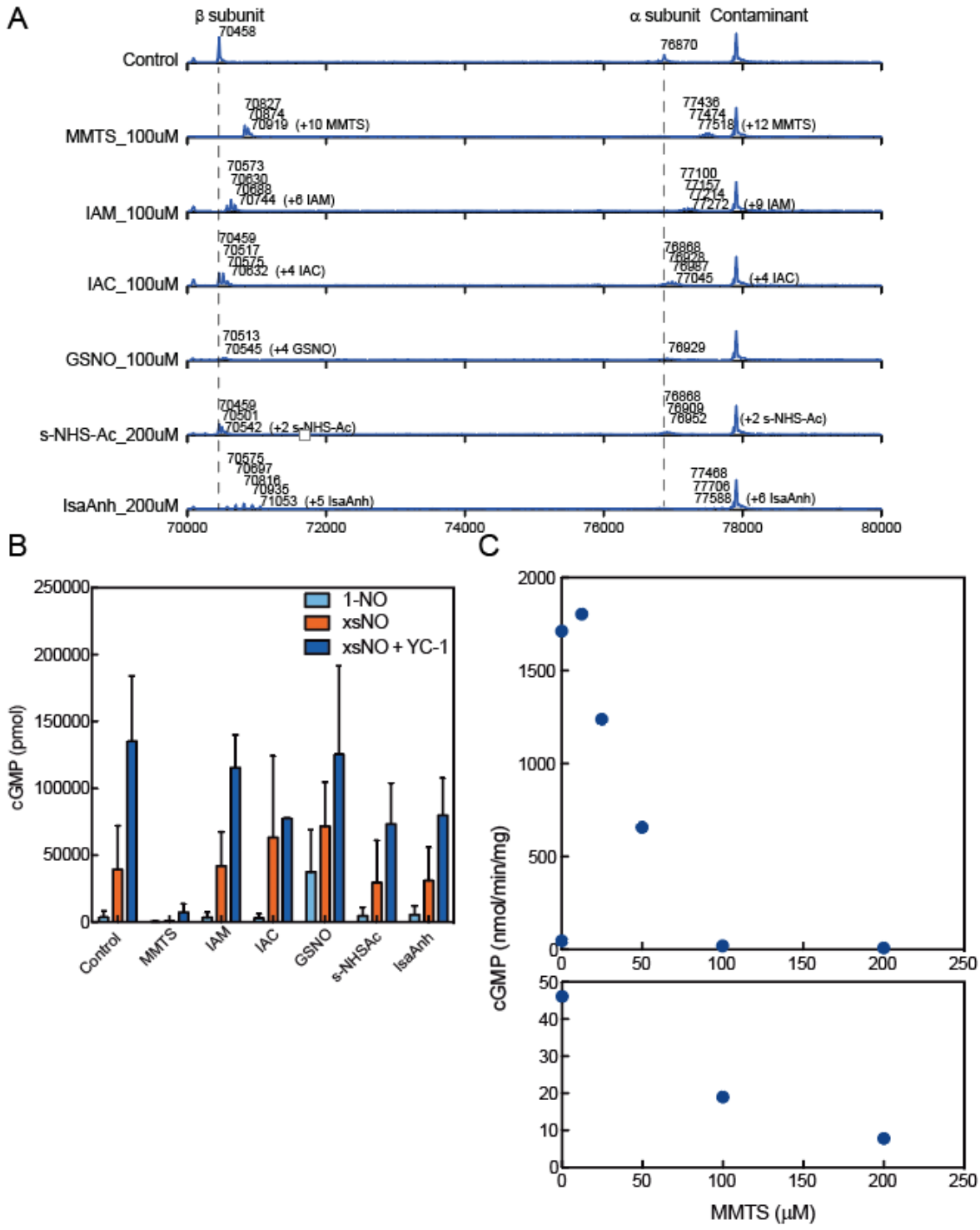


Figure 3.8 – A tandem mass spectrometry activity assay for sGC activation. A) Deconvoluted intact mass spectrometry of sGC with labeling conditions noted on the left. The masses of the varied modification states are noted, and the accompanying number of modifications is shown for the highest observed peak. B) Activity data with the same samples of protein that were examined by mass spectrometry. MMTS decreased excess NO-stimulated activity, however IAM and IAC did not in this assay. C) A titration of MMTS with 2 μM *Ms* sGC showing that MMTS between 50 and 100 μM brings the NO-stimulated activity back down to the unliganded activity state (noted by the low activity point at 0 μM MMTS).

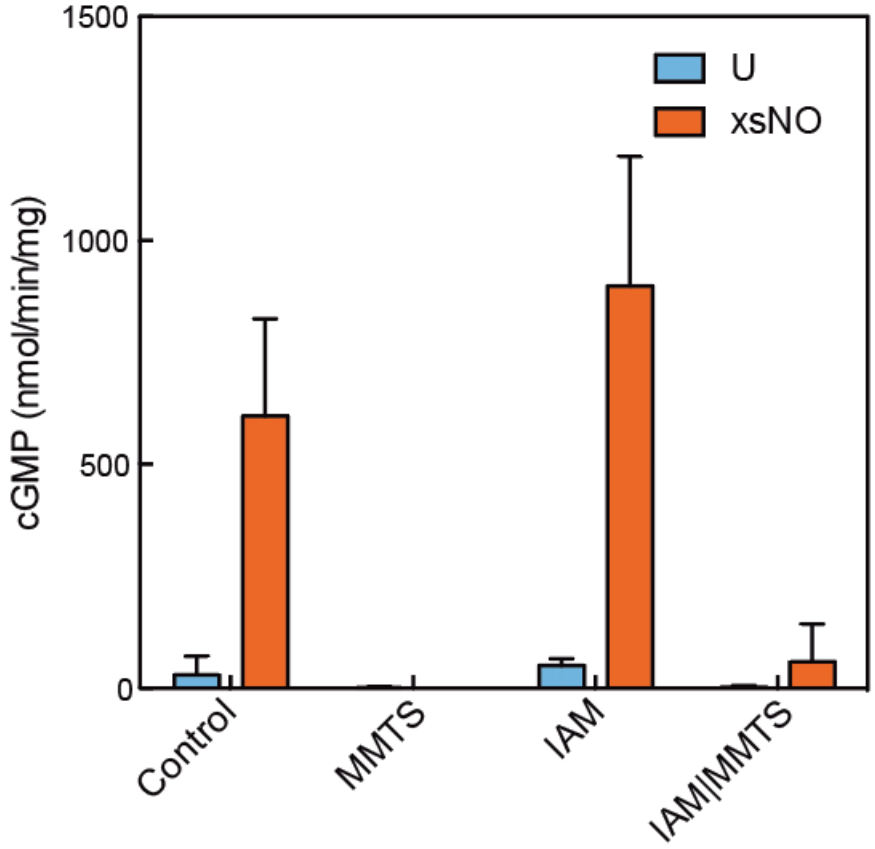


Figure 3.9 – Order of addition sGC activity assays. Endpoint activity data of *Ms* sGC showing that the reactivity of bioconjugational agent has an effect on NO-stimulated activity. In all three cases, the final concentration of alkylating agent was 100 µM, added 50 µM at a time with the first addition at 0 minutes and the second addition at 30 min. The IAM|MMTS sample had 50 µM of IAM added at 0 minutes, and 50 µM of MMTS added at 30 minutes.

Conclusions

A critical aspect of understanding molecular level function of signal transduction pathways involves locating the binding site of ligand(s) and elucidating how the signal is transduced from that location through the protein. Only then can detailed mechanistic understanding be applied to the root causes of disease states and ultimately develop therapeutics. The three different activation states of sGC are a fascinating biochemical phenomenon as the activation between 1-NO and excess NO is quite large, yet it is short lived and highly dependent on NO concentration. In this chapter, different sGC homologues and new evidence were applied to this problem to further our understanding of the three-state activation of sGC.

This investigation began with the hypothesis that the second site of NO binding is a cysteine residue. While new hypotheses were tested by examining conserved cysteines of other sGCs that have been biochemically confirmed to also possess a three-state activation profile, as well as other sGC homologues that cluster together in sequence space, none of the selected cysteines examined by themselves represented the sole second NO binding site. Advances in mass spectrometric techniques led to new experiments that tested both the reactivity of sGC cysteines

and the ability for NO to block the second site cysteine. However, this question turned out to be more complex than other questions answered with these techniques and no target cysteines emerged from this line of inquiry. Upon returning to biochemical activity assays, different hypotheses were explored to make sense of the set of results collected thus far.

Contrary to results previously published[13], YC-1 did not rescue MMTS treated sGC in this analysis. Upon expansion of the bioconjugational agents used, different cysteine labeling agents affected the ability of sGC to be stimulated by excess NO in vastly different ways. Both MMTS and NEM inhibited the ability for sGC to be stimulated, while IAM and IAC had little effect. These results highlight the critical nature of labeling conditions for sGC activity, specifically average number of sGC amino acids labeled, time of reaction, temperature of reaction, and which specific residues were modified. A variety of experiments can now be attempted to understand this three-state activation profile that is unique to sGC (described in Chapter 5).

References

(1) Moncada, S.; Palmer, R. M.; Higgs, E. A. Nitric Oxide: Physiology, Pathophysiology, and Pharmacology. *Pharmacol. Rev.* **1991**, *43* (2), 109–142.

(2) Arnold, W. P.; Mittal, C. K.; Katsuki, S.; Murad, F. Nitric Oxide Activates Guanylate Cyclase and Increases Guanosine 3':5'-Cyclic Monophosphate Levels in Various Tissue Preparations. *Proc. Natl. Acad. Sci. U. S. A.* **1977**, *74* (8), 3203–3207.

(3) Stone, J. R.; Marletta, M. A. Soluble Guanylate Cyclase from Bovine Lung: Activation with Nitric Oxide and Carbon Monoxide and Spectral Characterization of the Ferrous and Ferric States. *Biochemistry* **1994**, *33* (18), 5636–5640.

(4) Mergia, E.; Friebe, A.; Dangel, O.; Russwurm, M.; Koesling, D. Spare Guanylyl Cyclase NO Receptors Ensure High NO Sensitivity in the Vascular System. *J. Clin. Invest.* **2006**, *116* (6), 1731–1737.

(5) Batchelor, A. M.; Bartus, K.; Reynell, C.; Constantinou, S.; Halvey, E. J.; Held, K. F.; Dostmann, W. R.; Vernon, J.; Garthwaite, J. Exquisite Sensitivity to Subsecond, Picomolar Nitric Oxide Transients Conferred on Cells by Guanylyl Cyclase-Coupled Receptors. *Proc. Natl. Acad. Sci.* **2010**, *107* (51), 22060–22065.

(6) Zhao, Y.; Brandish, P.; Ballou, D.; Marletta, M. A. A Molecular Basis for Nitric Oxide Sensing by Soluble Guanylate Cyclase. *Proc. Natl. Acad. Sci. U. S. A.* **1999**, *96* (26), 14753–14758.

(7) Stone, J. R.; Marletta, M. A. Spectral and Kinetic Studies on the Activation of Soluble Guanylate Cyclase by Nitric Oxide. *Biochemistry* **1996**, *35* (4), 1093–1099.

(8) Russwurm, M.; Koesling, D. NO Activation of Guanylyl Cyclase. *EMBO J.* **2004**, *23* (22), 4443–4450.

(9) Cary, S. P. L.; Winger, J. A.; Marletta, M. A. Tonic and Acute Nitric Oxide Signaling through Soluble Guanylate Cyclase Is Mediated by Nonheme Nitric Oxide, ATP, and GTP. *Proc. Natl. Acad. Sci.* **2005**, *102* (37), 13064–13069.

(10) Derbyshire, E. R.; Marletta, M. A. Butyl Isocyanide as a Probe of the Activation Mechanism of Soluble Guanylate Cyclase: Investigating the Role of Non-Heme Nitric Oxide. *J. Biol. Chem.* **2007**, *282* (49), 35741–35748.

(11) Cary, S. P. L.; Winger, J. a.; Derbyshire, E. R.; Marletta, M. A. Nitric Oxide Signaling: No Longer Simply on or Off. *Trends Biochem. Sci.* **2006**, *31* (4), 231–239.

(12) Horst, B. G.; Marletta, M. A. Physiological Activation and Deactivation of Soluble Guanylate Cyclase. *Nitric Oxide* **2018**, *77*, 65–74.

(13) Fernhoff, N. B.; Derbyshire, E. R.; Marletta, M. A. A Nitric Oxide/Cysteine Interaction Mediates the Activation of Soluble Guanylate Cyclase. *Proc. Natl. Acad. Sci.* **2009**, *106* (51), 21602–21607.

(14) Hu, X.; Murata, L. B.; Weichsel, A.; Brailey, J. L.; Roberts, S. A.; Nighorn, A.; Montfort, W. R. Allostery in Recombinant Soluble Guanylyl Cyclase from Manduca Sexta. *J. Biol. Chem.* **2008**, *283* (30), 20968–20977.

(15) Brandish, P. E.; Buechler, W.; Marletta, M. A. Regeneration of the Ferrous Heme of Soluble Guanylate Cyclase from the Nitric Oxide Complex: Acceleration by Thiols and Oxyhemoglobin. *Biochemistry* **1998**, *37* (48), 16898–16907.

(16) Ward, C. C.; Kleinman, J. I.; Nomura, D. K. NHS-Esters As Versatile Reactivity-Based Probes for Mapping Proteome-Wide Ligandable Hotspots. *ACS Chem. Biol.* **2017**, *12* (6),

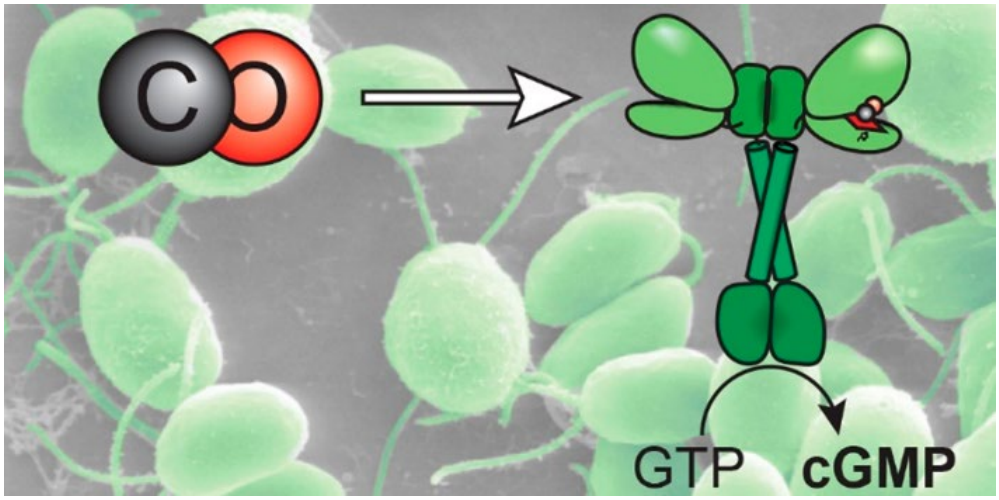
1478–1483.
(17) Underbakke, E. S.; Iavarone, A. T.; Chalmers, M. J.; Pascal, B. D.; Novick, S.; Griffin, P. R.; Marletta, M. A. Nitric Oxide-Induced Conformational Changes in Soluble Guanylate Cyclase. *Cell Struct.* **2014**, *22* (4), 602–611.
(18) Friebe, A.; Wedel, B.; Harteneck, C.; Foerster, J.; Schultz, G.; Koesling, D. Functions of Conserved Cysteines of Soluble Guanylyl Cyclase. *Biochemistry* **1997**, *36* (6), 1194–1198.
(19) Zallot, R.; Oberg, N.; Gerlt, J. A. The EFI Web Resource for Genomic Enzymology Tools: Leveraging Protein, Genome, and Metagenome Databases to Discover Novel Enzymes and Metabolic Pathways´. *Biochemistry* **2019**, *58*, 4169–4182.
(20) Crooks, G.; Hon, G.; Chandonia, J.; Brenner, S. WebLogo: A Sequence Logo Generator. *Genome Res* **2004**, *14*, 1188–1190.

CHAPTER 4

CHARACTERIZATION OF A CARBON MONOXIDE ACTIVATED SOLUBLE GUANYLATE CYCLAE FROM *CHLAMYDOMONAS REINHARDTII*

Summary

Signaling pathways that involve diatomic gases in photosynthetic organisms are not well understood. Exposure to nitric oxide or carbon monoxide is known to elicit certain responses in some photosynthetic organisms. For example, *Chlamydomonas reinhardtii* grown in low-iron media responds to exogenous carbon monoxide by increasing cell growth and intracellular chlorophyll levels. Here, we characterize Cyg11, a gas-responsive soluble guanylate cyclase from the eukaryotic green alga *Chlamydomonas reinhardtii* that converts GTP to cGMP. Cyg11 transcription is upregulated when *C. reinhardtii* is grown in iron-limited media, suggesting its importance in nutrient-limited environments. Cyg11 is purified as a homodimer and is activated by nitric oxide (2.5-fold over basal activity) and carbon monoxide (6.3 fold) The heme binding stoichiometry of Cyg11 was found to be one heme per homodimer, an unexpected result based on the sequence and oligomerization state of the enzyme. Gas binding properties, the kinetics of gas binding, and the ligand-modulated activity of Cyg11 are consistent with CO as the relevant physiological ligand.



Portions of this chapter have been reported in a previous publication.[1]

Introduction

Unicellular organisms must both sense and communicate with the outside world. Many of these communications occur through specific chemical interactions with selective receptor and signal transduction proteins, allowing the organisms to survive and proliferate in constantly-changing environments.[2,3] The origins of these interactions can be within or between species and are mediated through small molecules or even peptides;[4–10] however, signaling events can also be tuned to the surrounding environmental conditions.[11–13] In fact, the ability to rapidly sense environmental conditions is critically important for organisms to adapt to changes in temperature, pH, and availability of micronutrients. Understanding how unicellular organisms detect these external signals can provide insight into organism adaptation and the molecular mechanisms of signaling cascades. In turn, these systems can be used as a model to disentangle the more complex signaling systems in multicellular organisms.

Chlamydomonas reinhardtii, a eukaryotic unicellular green alga, is a well-suited model organism for the study of micronutrient sensing.[14] *C. reinhardtii* is known to respond to the changing availability of environmental micronutrients, such as copper, iron, and zinc. These environmental changes induce *C. reinhardtii* to degrade organelle-specific proteins, modulating both its transcriptome and proteome to maintain homeostasis.[15] For example, the nutritional copper sensing system of *C. reinhardtii* is particularly well understood. In low copper environments, both copper response proteins and a specific transcription factor, Copper Response Regulator 1 (CRR1), modulate the copper-dependent transcription of genes to preserve critical copper proteins.[16–18] This, and other examples, demonstrates the versatility of *C. reinhardtii* in determining the molecular mechanisms of signal transduction pathways.[19–21]

Diatomic gaseous signaling molecules present a unique challenge for aerobic organisms, as typical signaling gases these tend to be lower in concentration compared to oxygen (O_2). However, sensing systems that involve gaseous signaling molecules have been described in *C. reinhardtii*. The first such example demonstrates that nitrogen assimilation represses the transcription of nitrate reductase through a signaling system that potentially involves endogenously produced nitric oxide (NO).[22–24] Additionally, studies have linked exogenous carbon monoxide (CO) levels to the transcription of heme oxygenase and ferredoxin in low iron environments.[25,26] Furthermore, addition of CO to iron-depleted cells increases both the cell count and intracellular concentration of chlorophyll. These studies have implicated CO as a gaseous signaling molecule in *C. reinhardtii* in environments with low iron content.

The *C. reinhardtii* genome encodes six proteins homologous to single subunits of gas sensing soluble guanylate cyclases (sGCs) found in higher eukaryotes. In higher eukaryotes, sGCs are involved in many critical paracrine signaling processes, including vasodilation, platelet aggregation, and neurotransmission.[27–30] Higher eukaryotic sGCs are heterodimeric with each subunit consisting of four domains: an N-terminal Heme Nitric oxide/Oxygen (H-NOX) binding domain, a Per/Arnt/Sim (PAS)-like domain, a coiled-coil (CC) domain, and a catalytic cyclase domain (CAT). Typically, only one subunit of the heterodimer (usually denoted β) has the necessary residues to bind the heme cofactor. Based on the placement of the active site at the interface of the two subunit's catalytic domains and on sequence alignments of critical active site residues, most higher eukaryotic sGCs are necessarily heterodimeric and form one active site and

one pseudosymmetric site within their catalytic domain dimer.[31] sGCs catalyze the cyclization of 5′-guanosine triphosphate (GTP) to 3′,5′-cyclic guanosine monophosphate (cGMP). Enzyme activity is modulated by the presence of NO or O_2 at the heme binding site in the H-NOX domain of the β subunit.[32,33] cGMP in turn activates cGMP-dependent kinases, cGMP-gated ion channels, and other proteins initiating a diverse set of physiological responses. One *Cr* sGC, Cyg56, has been implicated in the repression of nitrate reductase signaling pathway by ammonium.[22] While the function of these proteins is well characterized in higher eukaryotes, the biochemical functions of the other sGCs in *C. reinhardtii* are largely unknown.

Reported here is a biochemical characterization of Cyg11 (Phytozome: Cre07.g320700.t1.1, UniProt A0A2K3DJ13), the first sGC with a higher maximum velocity of cGMP formation when CO is bound compared to when NO is bound. Additionally, Cyg11 was found to have a fast NO off-rate for a ferrous heme protein and a lower K_D for CO than any previously characterized sGC. These results suggest that Cyg11 plays a different role in signaling compared to sGCs in higher eukaryotes. Indeed, the results reported here provide support for CO as a signaling molecule in *C. reinhardtii* through the receptor Cyg11.

Materials and Methods

Materials

All chemicals were purchased from commercial vendors and used without further purification unless otherwise noted. Primers were obtained from Integrated DNA Technologies (Coralville, IA). Gibson Master Mix was purchased from New England Biolabs (Ipswich, MA). PrimeSTAR Max DNA polymerase and His60 Ni Superflow resin was purchased from Takara Bio USA (Mountain View, CA). DNA purification kits were purchased from Qiagen (Germantown, MD). Terrific Broth media powder, ampicillin, and 4-(2-aminoethyl)benzenesulfonyl fluoride hydrochloride (AEBSF) were purchased from Research Products International (Mount Prospect, IL). Hemin chloride, sodium dithionite ($Na_2S_2O_4$), DNAse I from bovine pancreas, β-mercaptoethanol, H_2NaPO_4, $NaCH_3CO_2$, $Mg(CH_3CO_2)_2$, and guanosine 5′-triphosphate sodium salt (GTP) (>95%, HPLC) were purchased from Millipore Sigma (Burlington, MA). Isopropyl β-D-1-thiogalactopyranoside (IPTG), 4-(2-Hydroxyethyl)piperazine-1-ethanesulfonic acid (HEPES), Pierce BCA Protein Assay Kit, ammonium acetate ($NH_4CH_3CO_2$), and acetonitrile (CH_3CN, HPLC grade) were purchased from Thermo Fisher Scientific (Waltham, MA). Benzamidine, lysozyme from chicken egg white, and sodium chloride (NaCl) were purchased from VWR Life Science (Visalia, CA). Dithiothreitol (DTT) was purchased from Bachem (Bubendorf, Switzerland). Imidazole was purchased from Oakwood Chemical (West Columbia, SC). Vivaspin spin concentrators were purchased from Sartorius (Concord, CA). PD-10 desalting columns were purchased from GE Healthcare (Pittsburgh, PA). DEA NONOate was purchased from Cayman Chemical Company (Ann Arbor, MI). Nitric oxide (NO, >99.5%) gas and carbon monoxide (CO, >99%) gas were purchased from Praxair Inc. (Danbury, CT). NO gas was bubbled through 10 M NaOH to remove higher oxides prior to use.

Construction of Expression Plasmids

The Core cDNA library of *Chlamydomonas reinhardtii* was purchased from the Chlamydomonas Resource Center (University of Minnesota). Primers were designed against the 3′ and 5′ UTR of Cyg11 (transcript name: Cre07.g320700.t1.1) Fwd: 5′-

GGATCCATCGATGCTTAGGAGGTC GCCATGCAGTTCATTCTTAGT -3′, Rev: 5′-CTGCTCATGTTTGACAGCTTATCATCG CGTGTCCGCGAGCAATGGTAA -3′. The underlined regions of the primers anneal to 3′ and 5′ UTR of Cyg11. The purified PCR product was subcloned into a linearized pCW vector with Gibson Master Mix, and the product was verified by DNA sequencing (UC Berkeley DNA Sequencing Facility). The Cyg11 predicted protein sequence was aligned to the β-subunit protein sequence of *Rattus norvegicus* (*Rn*) sGC (Uniprot: Q02153-1) using Clustal Omega (EMBL-EBI). The crystal structure of the *Homo sapiens* catalytic domain (99% identical to *Rn*, PDB ID: 4NI2) was used to identify the last ordered residue, and thus the N-terminal 615 amino acids were subcloned from the full-length Cyg11 gene. The final expression plasmid (pCW_Cyg11$_{615}$_His$_6$) includes a hexahistidine tag at the C-terminal of the protein, as the N-terminus of H-NOX domains are buried in the heme binding pocket and cannot be extended. The full peptide sequence is included in Figure B.1.

Protein Expression and Purification

pCW_Cyg11$_{615}$_His$_6$ was transformed into RP523(DE3) *E. coli* cells[34,35] and multiple colonies were tested for protein expression. After overnight growth in Terrific Broth (Research Products International) in the presence of 100 μg/mL ampicillin, 20 μg/mL hemin chloride (Millipore) (from a 10 mg/mL stock dissolved in DMSO), and 0.2% glucose, the cells were subcultured 1:100 into 1L and grown at 37 °C. A glycerol stock was created for each overnight culture. Once the cell density reached OD_{600} = 0.6, the culture temperature was lowered to 18 °C and protein expression was induced with 0.5 mM IPTG and grown for 16 hours. Expression levels were determined by pelleting approximately 10^8 cells pre- and post-induction, lysing with BugBuster (Millipore), running the cell lysate on a sodium dodecyl sulfate polyacrylamide (SDS-PAGE) gel, and performing a western blot to probe for His-tagged proteins using a mouse anti-His$_6$ Ab followed by a goat anti-mouse HRP conjugate Ab (Life Technologies). After the confirmation of an RP523 clone with high expression of Cyg11, overnight cultures were inoculated directly from the glycerol stock into liquid culture. Expression cultures (1 L) were spun down at 4300 *g* for 20 minutes, snap frozen in liquid nitrogen, and stored at –80 °C.

All following steps were performed at 4 °C. Cell pellets were thawed in ice water and resuspended in Lysis Buffer: Buffer A (50 mM Na_2HPO_4, pH 8.0, 200 mM NaCl, 1 mM imidazole, 5% (v/v) glycerol, 0.22 μm filtered) supplemented with 10 mM benzamidine, 1 mM AEBSF, 0.5 mg/mL lysozyme, 0.5 mg/mL Bovine DNAse I. Cells were lysed using a high-pressure homogenizer (Avestin Emulsiflex-C5). The resulting cell lysate was clarified by spinning at 27,000 *g* (Sorvall, SS-34 rotor) for 30 minutes, then gently rocked with 4 mL His60 resin, pre-equilibrated with Buffer A, for 30 minutes. The beads were applied to a column and washed with 10 CV of Buffer A. The column was washed with buffers with increasing concentrations of imidazole, and finally the protein was eluted with 5 CV of Buffer B (50 mM Na_2HPO_4, pH 8.0, 200 mM NaCl, 400 mM imidazole, 5% (v/v) glycerol, 0.22 μm filtered). The elution fractions were supplemented with 5 mM β-mercaptoethanol, concentrated to 5 mL using a 30,000 molecular weight cutoff spin concentrator, and passed over a 26/60 Superdex S200 pg size exclusion column (GE Healthcare) equilibrated with Buffer C (50 mM TEA, pH 7.5, 150 mM NaCl, 5 mM DTT, 5% glycerol, 0.22 μm filtered). Fractions that possessed a 430 nm/280 nm absorbance ratio greater than or equal to 1 (typically 1.2) were pooled, snap frozen in liquid nitrogen, and stored at –80 °C.

Analytical Size-Exclusion Chromatography

The following protein standards (0.2 mg/mL) were dissolved in Buffer C: lysozyme (14.3 kDa), deoxyribonuclease I (31 kDa), bovine serum albumin (66.5 kDa), alcohol dehydrogenase (150 kDa), and thyroglobulin (500 kDa). These standards and Cyg11 at 0.5 mg/mL were run over a Superdex 200 Increase 10/300 GL column (GE Healthcare) using Buffer C at a flow rate of 0.5 mL/min. The elution times were normalized to the void volume and plotted versus the log of the molecular weight to generate a standard curve.

Intact Protein Mass Spectrometry

Purified proteins were buffer-exchanged into 25 mM HEPES, pH 7.5, 25 mM NaCl buffer by three dilution and concentrations in a Vivaspin 500 (30,000 MWCO) spin concentrator (Sartorius). Samples were filtered by passing them through a 0.22 μM spin filter (Millipore). The final sample concentration was approximately 5 μM. Samples were separated with an Agilent 1200 series high pressure liquid chromatography (HPLC) system over a ProSwift column (ThermoFisher Scientific), and subsequently analyzed by an Agilent 6224 time-of-flight (TOF) mass spectrometer with a Turbospray ion source in positive ion mode.

Size-Exclusion Chromatography Multi-Angle Light Scattering (SEC-MALS)

Samples were injected onto a Superdex 200 10/300 SEC column (GE Healthcare) equilibrated with Buffer C at 0.5 mL/min. Elution peaks were detected with an in-line DAWN HELIOS II light scattering detector and an Optilab rEX refractive index detector (Wyatt Technologies). All steps were performed at room temperature. Data acquisition and processing were performed using Wyatt's ASTRA software program.

Native Nanoelectrospray Mass Spectrometry

Samples were prepared anaerobically in a glove bag (Coy) in an atmosphere of 98% Ar, 2% H. The sample was reduced with 5 mM $Na_2S_2O_4$ for fifteen minutes. Next, protein was buffer exchanged into 50 mM $Na_2CH_3CO_2$, pH 6.8, 2 mM DTT by diluting and concentrating the protein twenty times in a Vivaspin 500 (30,000 MWCO) spin concentrator (Sartorious) at 4 °C. The final protein concentration was 6 μM. The protein was removed from the anaerobic chamber and native nanoelectrospray ionization mass spectrometry (nanoESI-MS) was conducted on a Synapt G2-S*i* mass spectrometer equipped with a nanoESI source (Waters, Milford, MA) in positive ion mode. Mass spectrometry data acquisition and processing were performed using MassLynx software, version 4.1 (Waters).

Absorption Spectroscopy

Samples were reduced in an anaerobic glove bag (Coy) with 5 mM $Na_2S_2O_4$ for fifteen minutes, and then desalted with a PD-10 Minitrap desalting column (GE). CO-saturated buffer (950 μM CO) was prepared by sparging 3 mL of anaerobic Buffer D (50 mM HEPES, pH 7.5, 150 mM NaCl, 5% (v/v) glycerol, 0.22 μm filtered) for 15 minutes in a Reacti-Vial (Thermo Fisher Scientific). CO was added to the sample at a final concentration of 425 μM. Nitric oxide was added to a concentration of 500 μM by addition of DEA NONOate (Cayman Chemical Company) considering 1.5 moles of NO released for each mole of NONOate. Protein-gas complexes were incubated for 15 minutes at room temperature to establish equilibrium, and no further spectral changes were observed after this time. Samples were placed in a septum-sealed 1 cm pathlength quartz cuvette inside the glove bag, and UV–Vis spectra were recorded on a Cary 300 spectrophotometer (Agilent Technologies).

Heme Reconstitution
Samples were reduced in an anaerobic glove bag (Coy) with 5 mM DTT, and then 2 equivalents to homodimer of hemin chloride in 100% DMSO (Millipore Sigma) was added (final DMSO concentration was 1%). No difference in reconstitution was observed when samples were incubated for 15 minutes at 25 °C or 16 hours at 4 °C. Reconstituted samples were desalted over PD-10 Minitrap desalting column (GE), and absorbance spectra were recorded as above.

Activity Assays and Quantification
Steady-state kinetics for Cyg11 were measured by quantifying the amount of cGMP produced in endpoint activity assays. Protein-gas complexes were prepared as described for absorption spectroscopy. For all steady-state kinetic assays, the final protein concentration was approximately 1 μM, as quantified by Nanodrop A_{280} measurement (molecular weight of Cyg11 homodimer is 139,429 Da, extinction coefficient calculated using ExPasy ProtParam to be 52,830 $M^{-1}cm^{-1}$ assuming all cysteines are reduced). Activity assays were conducted at 25 °C and pH 7.5 in Buffer D supplemented with 5 mM DTT and $Mg(CH_3COO_2)_2$ greater than or equal to twice the substrate concentration. Reactions were initiated with GTP and timepoints were quenched into 1% formic acid. Reactions were analyzed by reverse-phase HPLC to quantify the amount of cGMP formed. Samples (88 μL) were injected onto a pre-equilibrated Eclipse Plus C18 column, 3.5 μm particle size (4.6 x 100 mm) (Agilent). The HPLC system used two buffers (Buffer E: 20 mM ammonium acetate, 0.1% v/v formic acid; Buffer F: 99.9% acetonitrile, 0.1% formic acid) and was run at 1 mL/min to pre-equilibrate with 2% Buffer F. After sample injection, the following gradient was run at 1 mL/min to separate substrate from product: 2% Buffer F for minutes 0–6, 2–25% Buffer F for minutes 6–7.5, 25–100% Buffer F for minutes 7.5–8.5, 100% Buffer F for minutes 8.5–9, 100–2% Buffer F from minutes 9–9.25, and 2% Buffer F for minutes 9.25–13.5. The peaks of GTP (1.7 min) and cGMP (4.2 min) were monitored at 254 nm. Concentrations of cGMP were determined from a standard curve, generated over 6–200 μM. Initial rates were calculated from the linear phase of the time course during which 5–10% of the GTP substrate was consumed.

Gas Binding Kinetics
Gas dissociation rates (k_{off}) were determined for CO-bound and NO-bound Cyg11. The protein was reduced anaerobically and desalted as before and either diluted in CO-saturated buffer or DEA NONOate was added to give an initial concentration of 5 μM protein complex. The trap solution for CO dissociation rates was prepared by sparging $NO_{(g)}$ into Buffer H (50 mM HEPES, pH 7.5, 150 mM NaCl, 0.22 μm filtered) for 15 minutes in a Reacti-Vial. The trap solution for NO dissociation rates was prepared by adding $Na_2S_2O_4$ to Buffer H and sparging with $CO_{(g)}$ for 15 minutes in a Reacti-vial. Samples were transferred in gas-tight syringes to a HiTech KinetAsyst (TgK Scientific) stopped-flow system. Reactions were initiated by mixing an equal volume of Cyg11–CO with $NO_{(g)}$ saturated buffer trap, or Cyg11–NO with CO/dithionite buffer trap at 20 °C. The concentration of the $NO_{(g)}$ (95–950 μM) and dithionite (10–100 mM) was varied to ensure that the rate was independent of trap concentration.

CO association rates (k_{on}) were determined for ferrous Cyg11. Samples (approximately 5 μM) were transferred in gas-tight syringes into the stopped-flow system that was made anaerobic by incubating the sample-handling unit with 10 mM $Na_2S_2O_4$ for 30 minutes. Dilutions of CO-saturated buffer were made on the sample handling unit in gas-tight syringes by diluting CO-

saturated buffer with anaerobic Buffer H via a three-way valve. Reactions were initiated by mixing equal volumes of protein solution and CO-saturated buffer at 20 °C for a final protein concentration of 2.5 μM.

Results and Discussion

Identification and sequence analysis of C. reinhardtii sGCs

Sequence alignments of all six annotated sGC homologues from *Chlamydomonas reinhardtii* (*Cr*) were compared to the mammalian sGC from *Rattus norvegicus* (*Rn*) (Figure 4.1).[22] All *Cr* sGCs have four predicted domains similar to mammalian sGC. A schematic of the domain organization is shown for both Cyg11 and *Rn* sGC $\alpha_1\beta_1$ in Figure 4.1A. All *Cr* sGCs are predicted to contain a long C-terminal domain with no homology to any known protein domain. It is not known what function these C-terminal domains have in *Cr* sGCs, or if these domains are present in the mature proteins *in vivo*.

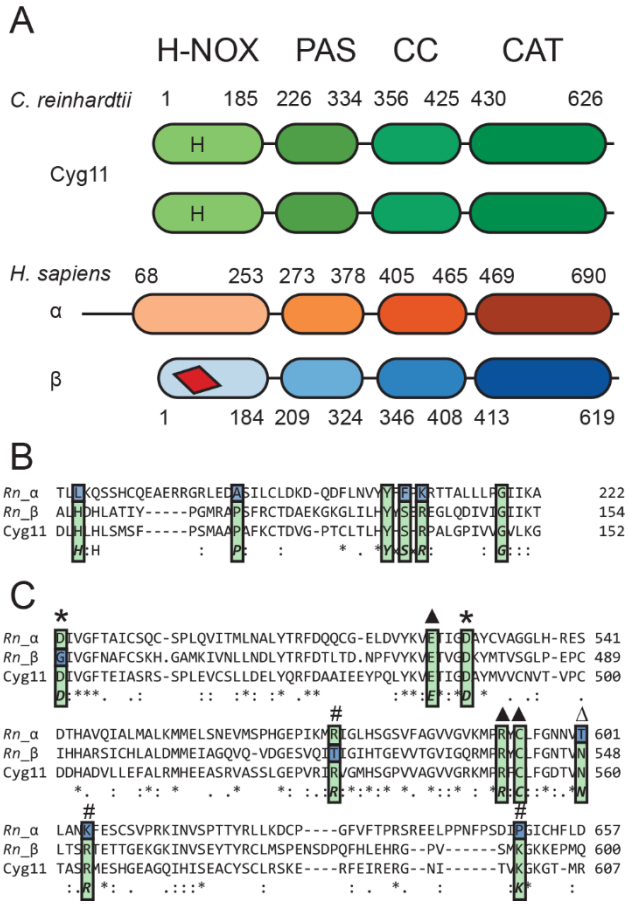


Figure 4.1 – Domain arrangement of *Cr* sGCs. As homologues to mammalian sGC, *Cr* sGCs are composed of four consecutive domains: a Heme Nitric Oxide/Oxygen binding domain (H-NOX), a Per/Arnt/Sim domain (PAS), a coiled-coil domain (CC), and finally the catalytic domain (CAT). A) A cartoon schematic of the domain organization and approximate domain regions of both Cyg11 and *Rn* $\alpha_1\beta_1$ sGC. The letter H represents possible heme binding sites in each Cyg11 subunit. B) Sequence alignments of the H-NOX domain of *Cr* Cyg11 and *Rn* sGC. All *Cr* sGC H-NOX domains have a proximal histidine, and the YxSxR motif necessary to bind a heme cofactor (see Figure B.1 for other *Cr* sGC sequences) C) Sequence alignments of *Cr* Cyg11 and *Rn* sGC CAT domains. Whereas the heterodimeric structure of *Rn* sGC is required to generate a competent active site, *Cr* sGC CAT domains contain all the residues

necessary for catalysis, rendering them catically-active homodimers. Key residues: * – metal binding, Δ – guanine binding, ▲ – triphosphate binding, # – ribose binding.

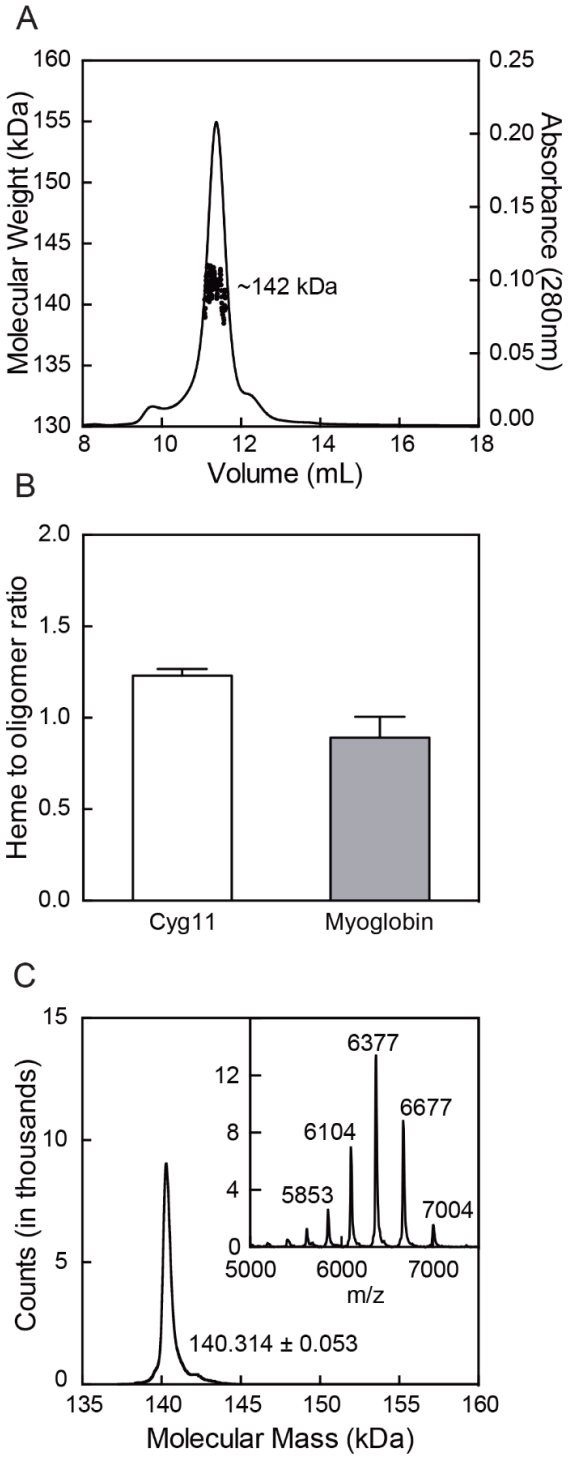
Rn sGC is a heterodimer composed of an α_1 and a β_1 subunit. Only the β_1 H-NOX domain contains the YxSxR motif, which binds the propionate groups of the heme cofactor and a histidine residue to act as the proximal ligand to the iron. *Cr* sGCs showed more homology to the heme-binding *Rn* β_1 sGC than *Rn* α_1 (average of 36% versus 28% identity), including all residues mentioned above to bind heme in the H-NOX domain (Figure 4.1B, Figure B.1B). As all *Cr* sGC H-NOX domains have the heme-binding motif, it is not possible to distinguish between gas-sensing homodimers versus heterodimers as is the case for the *Rn* sGC.

Interestingly, in contrast to the *Rn* $\alpha_1\beta_1$ sGC that must form a heterodimer to compose one complete active site and one pseudosymmetric site, *Cr* sGC CAT domains contain all necessary residues to form two competent active sites. As such, the *Cr* sGCs are predicted to form both active homodimers and active heterodimers (Figure 4.1C, Figure B.1B). Similar to sequence analysis of the H-NOX domain, these findings do not aid in distinguishing which oligomers are formed *in vivo*, so transcriptomics was used to narrow down the possibilities.

Analysis of Trancriptomics Data

Available transcriptomic data for *Cr* sGCs was obtained from Phytozome (Version 12) in order to help discern the oligomerization state of *Cr* sGCs.[36] All six predicted homologues of mammalian sGC from *C. reinhardtii* are transcribed at low, but detectable levels (Figure B.2). Co-transcriptional analyses across all studies indicated that no two *Cr* sGCs are predicted to be transcribed together greater than 85% of the time. While transcriptomic data does not guarantee protein expression as would be confirmed by proteomic data, the fact that none of the six sGCs are predicted to co-transcribed supports the hypothesis that *Cr* sGCs form homodimers. It is uncommon for higher eukaryotic sGCs to form homodimers natively, however this greatly facilitates heterologous expression and further characterization for *C. reinhardtii* sGCs.

Since transcriptomics data for varying gas conditions are not available, conditions that may be important in gaseous signaling were analyzed. Under iron-deficient conditions, exogenous CO has been shown to increase cell growth, increase cellular chlorophyll concentrations, and upregulate genes involved in iron metabolism.[25,26] Consequently, the transcripts for *C. reinhardtii* genes with varying levels of iron were analyzed (either replete, deficient, or limited).[37] The transcript levels of two *Cr* sGCs (*cyg11* and *cyg15*) increase as iron levels decrease, indicating that both genes are upregulated in low iron environments. Furthermore, the absolute number of *cyg11* transcripts increases when *C. reinhardtii* is grown photosynthetically, potentially indicating the importance of this gene in a native environment. Since Cyg11 is transcribed at higher levels under normal physiological conditions, this enzyme was selected for further study to determine if it may be involved in CO-mediated recovery from iron deficiency.[25,26]



Initial Characterization of Cyg11

Cyg11 was expressed heterologously in *E. coli* RP523(DE3), a heme-permeable strain, to ensure excess heme is present during protein expression.[34] Cyg11 was purified by immobilized metal affinity chromatography followed by size-exclusion chromatography (Figure B.3,B.4). The protein was isolated with the heme in the five-coordinate ferrous state. To confirm Cyg11 formed homodimers, size exclusion chromatography multiangle light scattering (SEC-MALS) experiments were performed with the instrument calibrated to a BSA standard. As shown in Figure 2A, Cyg11 eluted in a single peak (V_r = 11.3 mL) with a calculated molecular weight of 142 kDa, which is nearly double the expected molecular weight (69.7 kDa, Figure B.5). This result clearly shows that Cyg11 forms homodimers.

The Cyg11 homodimer has two H-NOX domains; thus, the expectation was that each homodimer would bind two heme cofactors. The heme stoichiometry was measured using two assays performed in tandem: a pyridine hemochromogen assay to measure heme concentration and a bicinchoninic acid (BCA) assay to measure protein concentration.[38] The results of this tandem assay indicated that there was approximately one heme cofactor per dimer of Cyg11 (Figure 4.2B). Native mass spectrometry of the purified Cyg11 showed one homogeneous protein population with regards to heme occupancy (Figure 4.2C, inset). The deconvoluted mass spectrum of Cyg11 revealed a single mass of 140,314 ± 53 Da. This mass confirms that Cyg11 does not contain two heme cofactors, as a dimer with two hemes would have a mass of 140,660 Da (Figure 4.2C). A homodimer with one heme is 140,044 Da, within range of the measured mass plus bound cations from the buffer. Attempts to reconstitute Cyg11 with exogenous hemin

Figure 4.2 – Oligomeric state and heme occupancy of Cyg11 A) Size-exclusion multi-angle light scattering trace shows Cyg11 eluting in one main peak with a molecular weight around 142 kDa, about twice the monomer weight of 70 kDa. B) A pyridine hemochromogen assay, coupled with a bicinchoninic acid (BCA) assay, suggests that Cyg11 has only one heme per dimer, using myoglobin with one heme as a reference. C) Raw (inset), and deconvoluted native nano-electrospray mass spectrometry confirms the dimeric state of Cyg11 with only one heme cofactor bound to the purified enzyme (a dimer with two heme cofactors is 140,660 Da).

slightly increased the ratio of the Soret band to protein absorbance at 280 nm but did not increase the content to two heme cofactors per dimer (Figure B.6). Both the tandem assay and mass spectrometry indicate that Cyg11 contains one heme per homodimer, an unexpected result for a homodimer with two potentially functional heme binding domains. It is not known when heme loading occurs in the maturation of Cyg11. This final occupancy could be the result of homodimer formation that creates one functional heme binding site or loss of one heme upon dimerization.

Diatomic Gas Binding to Cyg11

All known sGCs are activated with diatomic gases. Thus, to assess which ligands the Cyg11 H-NOX domain binds, UV-visible absorption spectroscopy was used to determine ligand-dependent shifts in spectra of the heme cofactor. Cyg11 purifies in the ferrous state, with a λ_{max} of the Soret band centered at 431 nm, consistent with a 5-coordinate ferrous heme as isolated. Cyg11 readily binds both NO and CO, resulting in blue shifts to 399 nm (5-coordinate complex) and 424 nm (6-coordinate complex), respectively (Figure 4.3). The α/β bands also show the expected increased splitting upon the addition of these diatomic gases (Figure 4.3, inset). Upon exposure of Cyg11 to atmospheric oxygen, no change in the Soret band was observed (Figure B.7). That Cyg11 does not bind oxygen is consistent with the sequence of its H-NOX domain, which lacks the critical hydrogen bonding residues required for O_2 binding.[39] Taken together, the ligand-dependent absorption spectra for Cyg11 is similar to that of both mammalian sGCs and *Shewanella oneidensis* (*So*) H-NOX, which indicates that Cyg11 likely senses either NO or CO *in vivo* (Table B.1).[40,41]

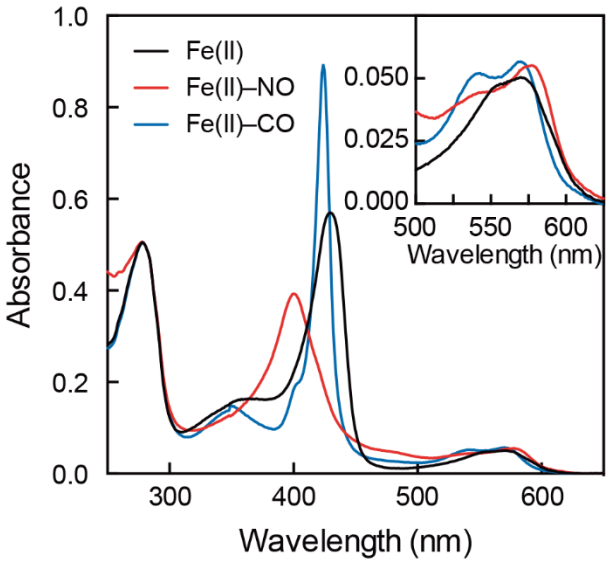
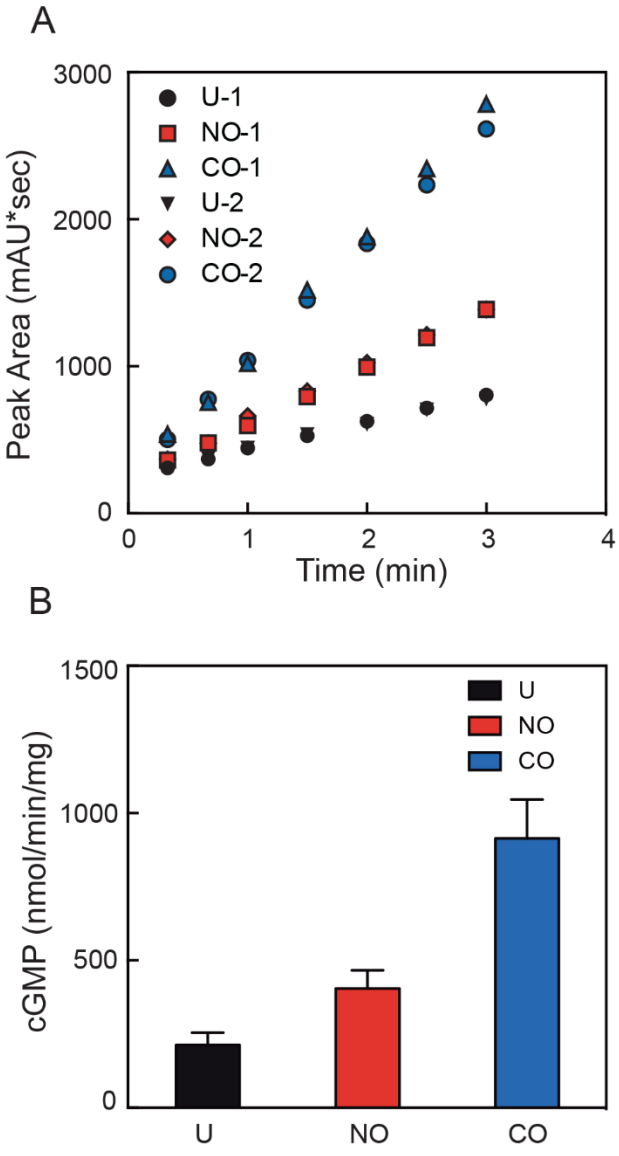


Figure 4.3 – UV–visible absorption spectra of Cyg11 The UV–visible absorption spectra of Cyg11 with various gaseous ligands. The Soret band of Cyg11 is sensitive to the identity of the axial ligand, blue-shifting for CO and more so for NO. Inset: The α/β exhibit increased splitting as the diatomic gases bind to the heme.

Kinetics of cGMP Formation by Cyg11

To discern the effects of ligand binding on enzyme activity, the steady-state kinetics of Cyg11 were measured with three different ligation states: unliganded, NO-bound, and CO-bound (Figure 4.4). Cyg11 displays a basal activity of 213 ± 42 nmol min^{-1} mg^{-1} at 2 mM GTP, which is higher than previously characterized sGC activities in the unliganded state[42]. Upon NO binding, the activity increases to 405 ± 62 nmol min^{-1} mg^{-1}. However, CO binding results in an increase to 914 ± 132 nmol min^{-1} mg^{-1}. Activity assays with Cyg11 reconstituted with hemin showed no difference when compared to Cyg11 not reconstituted with hemin (Figure B.8). While adding CO to previously characterized sGCs increases activity over basal, this is the first example of an sGC in which the CO-bound enzyme has a higher activity than the NO-bound enzyme.



A Michaelis-Menten curve was constructed for all three ligation states (Figure B.9) from 1-32 mM GTP. Product inhibition was observed before enzyme saturation is achieved, precluding the precise assignment of kinetic parameters. Estimates for k_{cat}/K_M were made by fitting the data up to and including 16 mM GTP to Equation 1:

$$\frac{V}{E} = \frac{\frac{k_{cat}}{K_M} \times \frac{[S]}{K_M}}{\left(1 + \frac{[S]}{K_M}\right)} \quad (1)$$

thereby placing greater significance on lower GTP concentrations. A summary of the results is provided in Table 4.1. CO-bound Cyg11 is significantly more active than the other species. In all cases, $K_{M,GTP}$ is estimated to be above physiological GTP concentrations (0.9–1.5 mM)[43,44]; this signifies the activity of Cyg11 will vary greatly with small changes in GTP concentration.

Figure 4.4 – Activity assay of Cyg11 with different ligation states A) Initial rates were taken from a discontinuous enzyme assay run at 25 °C and pH 7.5, and cGMP was measured using an HPLC assay. Two representative trials are shown. B) The initial rates are plotted from an assay with 2 mM GTP as the substrate. The average is plotted with the standard deviation as error bars from 5 independent experiments.

Table 4.1 – Cyg11 steady-state kinetic parameters

Ligation State	V_{max}/K_M (nmol/(min*mg))/mM
Unliganded	63.8 ± 6.20
NO	162.8 ± 39.70
CO	406.2 ± 64.96

Gas Binding Kinetics

The on- and off-rates of gas ligands provide insight into how long the signal is transduced. To garner a comprehensive view of the gas binding kinetics for Cyg11, CO and NO gas dissociation rates were measured using stopped-flow UV-vis spectroscopy.

The dissociation rate of CO from Cyg11 was measured by rapidly mixing of Cyg11–CO with $NO_{(g)}$-saturated buffer. This leads to the decrease of the Soret peak at 424 nm and the increase of the 5-coordinate ferrous nitrosyl peak at 399 nm (Figure 4.5A). The absorbance values of the difference spectrum were extracted (Figure 4.5A, inset) and fit to a monoexponential decay process with a rate constant $k = 6.04 \pm 0.22\ s^{-1}$, which is consistent with the CO dissociation rate of other H-NOX domains (Figure 4.5B, Table 2).

The converse experiment was also conducted, in which the rate of dissociation of NO from Cyg11 was measured by mixing Cyg11–NO with CO-saturated buffer containing a NO trap (30 mM $Na_2S_2O_4$). The Soret band maximum at 400 nm for the NO complex rapidly shifts to 424 nm with the formation of the CO adduct (Figure 4.5C). The NO dissociation rate did not depend on the concentration of $Na_2S_2O_4$ in the trap solution (10–100 mM). Dissociation rates were calculated using the change in absorbance at 424 nm (Figure 4.5D). The data were fit to both a one- (red) and a two-component (blue) exponential function. The NO dissociation process is best described with a double exponential function with rate constants of $k_1 = 0.13 \pm 0.02\ s^{-1}$ (67%) and $k_2 = 0.019 \pm 0.04\ s^{-1}$ (33%). Biphasic NO dissociation is typical for H-NOX domains and has been attributed to two electronically similar "open" and "closed" states of the protein.[45–49] Typically, the slow component dominates, and thus complete dissociation takes several hours to occur.[45] However, as seen in Figure 4.5D, this process is complete within 5 minutes for Cyg11, and the dominant component of the rate constant is three orders of magnitude greater than the dominant component of rate constants other characterized H-NOX domains (Table 2). This indicates that NO dissociates far faster from Cyg11 compared to other sGCs.

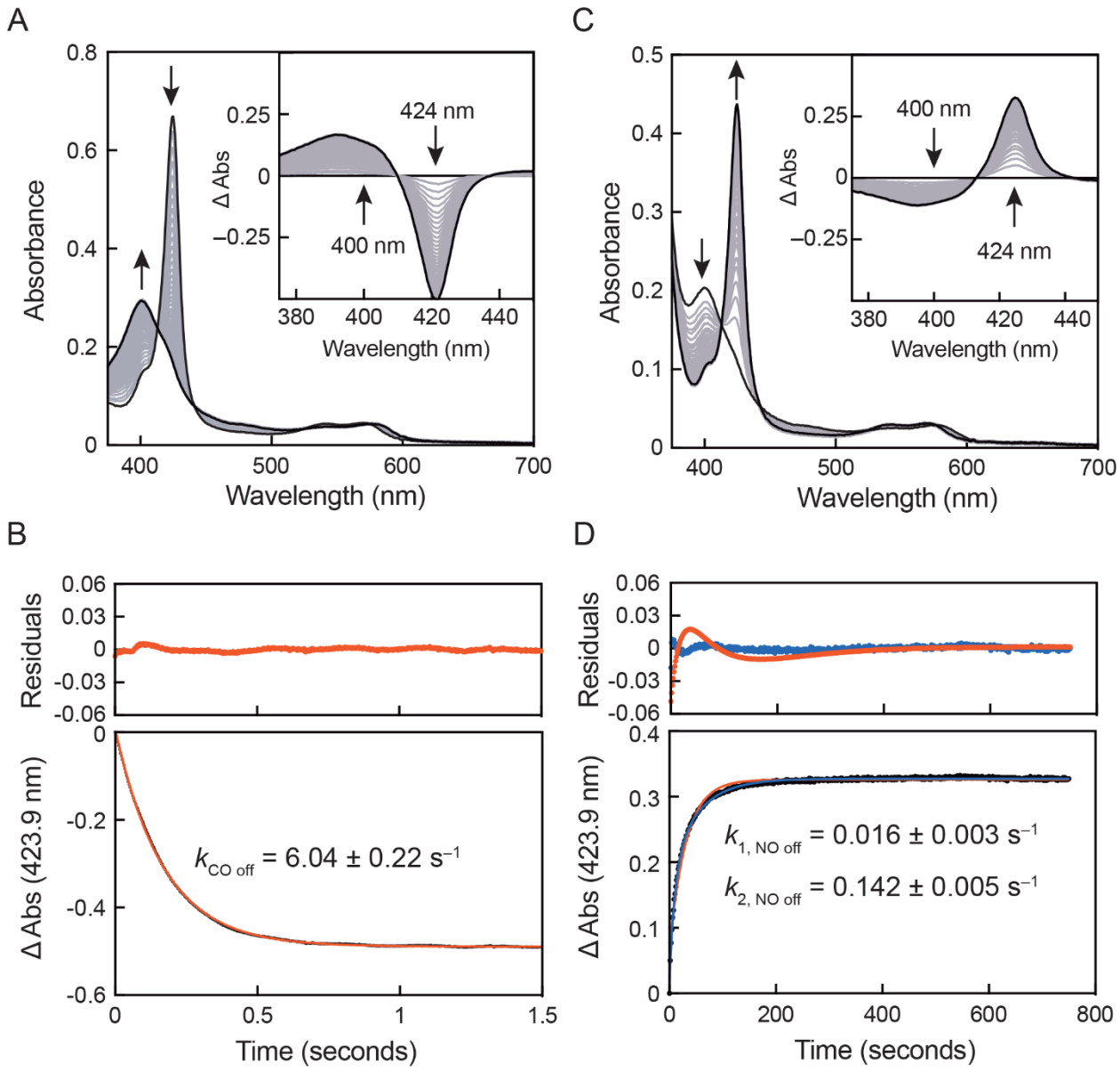
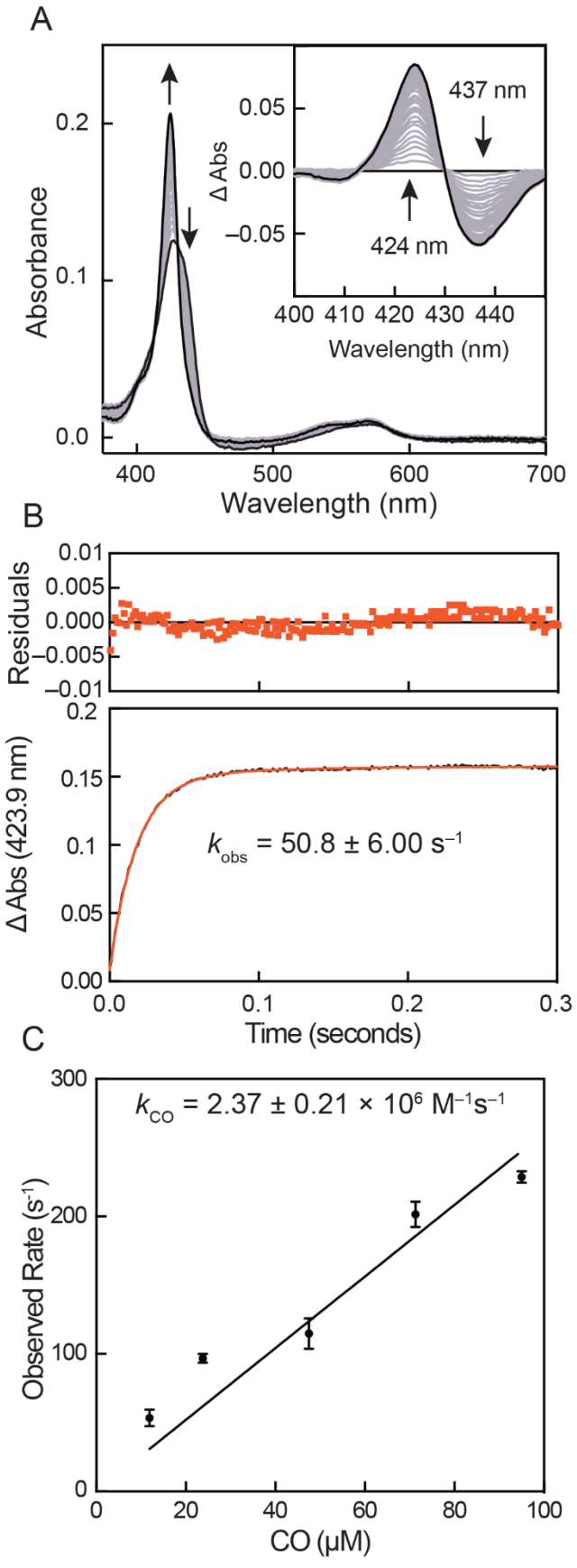


Figure 4.5 – Dissociation rates of CO and NO from Cyg11. The dissociation rates of CO and NO were measured using stopped-flow UV-visible absorption spectroscopy. A) Representative absorption spectrum for the dissociation of CO using excess NO as the trapping agent (inset: difference spectrum). Spectra were recorded every 0.003 seconds, and every fourth spectrum is shown for clarity. B) The change in absorbance at 424 nm was plotted as a function of time and fit to a monoexponential function. The residuals are shown above the plot. C) Representative absorption spectrum for the dissociation of NO using CO and $Na_2S_2O_4$ as a trapping agent (inset: difference spectrum). The concentration of $Na_2S_2O_4$ did not alter the dissociation rate. The spectra were recorded every 1.5 seconds, and every fourth spectrum is shown for clarity. D) The change in absorbance at 424 nm was plotted as a function of time, fitting the data to both a single (red) and double (blue) exponential function. The residuals are shown above the plot, indicating that a biexponential function best captures the experimental data.

CO association rates were measured by mixing equal volumes of unliganded Cyg11 with CO-saturated buffer. The change in absorbance was measured at 424 nm as a function of time, and the difference spectrum was fit to a single exponential function to obtain an observed rate constant, k_{obs} (Figure 4.6A and 6B). These experiments were repeated with varied CO concentrations (12-



95 μM) under pseudo-first order reaction conditions to measure the bimolecular rate constant for CO association. The observed CO association rates were then plotted as a function of CO concentration (Figure 4.6C) to give a rate constant of $k_{on,CO}$ = 2.37 ± 0.21 × 10^6 M^{-1} s^{-1}, which is three orders of magnitude greater than previously characterized full-length sGCs (Table 2). In conjunction with the CO dissociation rates measured above, dissociation constant for CO for Cyg11 can then be calculated to be $K_{D, CO}$ = 2.5 × 10^{-6} ± 0.24 × 10^{-6} M. This value is approximately the same as for *So* H-NOX ($K_{D, CO}$ = 1 × 10^{-6})[48] and two orders of magnitude smaller than the dissociation constant for the *Rn* sGC ($K_{D, CO}$ = 2.6 × 10^{-4} M).[50] As was seen for activation parameters, gas binding kinetics for Cyg11 deviate from previously-characterized sGCs.

The ability to discriminate between diatomic gases is a hallmark of gas-sensing proteins and H-NOX domains. NO-sensing H-NOX proteins use the femtomolar K_D of NO for ferrous heme and a hydrophobic gas-binding pocket to impart specificity. Conversely, O_2-sensing H-NOX proteins possess a hydrogen-bonding motif in the gas-binding pocket that stabilizes the heme O_2 adduct. Cyg11 has a K_D for CO that suggests it may be the physiological ligand.

Figure 4.6 – Cyg11 CO association rate. The association rate of CO to Cyg11 was measured using stopped-flow spectroscopy. A) Representative absorption spectrum for the association of CO to unliganded Cyg11 (inset: difference spectrum). Spectra were recorded every 0.003 seconds; every fourth spectrum is shown for clarity. B) The change in absorbance at 424 nm was plotted as a function of time. The data was fit to a single exponential decay and the residuals are shown above the plot. C) The CO concentration was varied in the trapping solution and the observed rates are plotted against the concentrations to give the association rate. The error bars represent the standard deviation of at least three measurements.

The dissociation rate constant for CO is in line with other H-NOX domains; however, the on-rate for CO is significantly faster than any other sGC or bacterial H-NOX protein. Accordingly, Cyg11 has the tightest affinity for CO of characterized full length sGCs. Furthermore, the kinetics of NO dissociation are significantly faster than typical sGCs, as the dominant rate constant for Cyg11 is three orders of magnitude faster than previously reported sGCs. While the K_D of NO is predicted to be much smaller than that of CO, we hypothesize that Cyg11 is transcribed and expressed under conditions when *C. reinhardtii* is not exposed to NO, thus eliminating competition for the heme binding pocket. Taken together, these results support the hypothesis that Cyg11 is acting as a putative CO sensing protein in *C. reinhardtii*.

Table 4.2 –Dissociation and Association Rates of CO and NO

	protein	k_1	k_2	ΔA1 (%)	ΔA2 (%)	temp
CO dis.	Cyg11	6.04 (0.22) s^{-1}				20 °C
	Rn sGC	10.7 s^{-1}				24 °C
	So H-NOX	0.85 s^{-1}				25 °C
NO dis.[b]	Cyg11	0.142 (0.005) s^{-1}	0.016 (0.003) s^{-1}	67 (14)	33 (9)	20 °C
	Rn sGC	0.0118 (0.0071) s^{-1}	0.00012 (0.00002) s^{-1}	5 (1)	95 (1)	10 °C
	So H-NOX	0.0014 (0.0001) s^{-1}	0.00028 (0.00001) s^{-1}	29 (6)	71 (6)	20 °C
CO as.	Cyg11	237 (21) × 10^4 M^{-1} s^{-1}				20 °C
	Rn sGC	4 × 10^4 M^{-1} s^{-1}				24 °C
	So H-NOX	82 × 10^4 M^{-1} s^{-1}				25 °C

[a]All measured and reported standard deviations are depicted in parentheses. [b]TW – This work.

Conclusion

All signaling systems require receptors to transduce a primary signal to secondary signals that interact with downstream proteins to elicit a response. Although a number of physiological responses to gas molecules have been described for plants and other photosynthetic organisms,[22,25,26,51–53] the biochemical characterization of gas receptors in these systems is in its infancy.[54] In this study, the eukaryotic alga *Chlamydomonas reinhardtii* is used as a model to understand gaseous signaling in response to micronutrient changes. This organism is adept at adapting to various micronutrient conditions in its environment; therefore, the organism must have discriminating sensing mechanisms to determine the concentrations of analytes in the surrounding areas and effectively respond to these stimuli.

The soluble guanylate cyclase Cyg11 from *Chlamydomonas reinhardtii* has some unique features compared to higher eukaryotic sGCs. First, it has some unusual and unexpected biochemical properties. It is active as a homodimer, so unique in the sGC family in that regard. Further, it has two H-NOX domains each competent to bind heme, and yet only one heme is found in the homodimer. Although consistent with one heme in higher eukaryote sGCs, this was an unexpected result. Second, the majority of those previously characterized sGCs demonstrate a decided preference for NO as both a ligand and activator of activity. The only exceptions to this are sGCs from *D. melanogaster*, whose activity is influenced by O_2 heme binding. As opposed to all other characterized sGCs, Cyg11 has a higher activity with CO rather than NO. Moreover, the enzyme binds CO tighter and has a faster NO off-rate than any previously reported sGC. In combination with transcriptomics data, the high affinity of Cyg11 for CO and the higher activity of Cyg11 with CO bound could indicate that it is a CO-sensing enzyme that *C. reinhardtii* uses to amplify the signal of low-iron.

From these results, we hypothesize that this enzyme is a putative CO sensor and may be involved in the CO-dependent recovery pathway that is activated under iron-deficient growth conditions. Furthermore, it is possible that Cyg11 is sensing CO is formed endogenously by the enzyme heme oxygenase-1 when heme is degraded from other non-essential enzymes. Since this is the first sGC homologue showing a preference for CO, cGMP produced by Cyg11 may be implicated as a secondary signaling molecule that enables *C. reinhardtii* to adapt to iron-deficient conditions.

References

(1) Horst, B. G.; Stewart, E. M.; Nazarian, A. A.; Marletta, M. A. Characterization of a Carbon Monoxide-Activated Soluble Guanylate Cyclase from Chlamydomonas Reinhardtii. *Biochemistry* **2019**, *58* (17), 2250–2259.

(2) Libert, S.; Pletcher, S. D. Modulation of Longevity by Environmental Sensing. *Cell* **2007**, *131* (7), 1231–1234.

(3) Ryan, R. P.; Dow, J. M. Diffusible Signals and Interspecies Communication in Bacteria. *Microbiology* **2008**, *154* (7), 1845–1858.

(4) Greenberg, E. P.; Fuqua, W. C.; Winans, S. C. Quorum Sensing in Bacteria : The LuxR-LuxI Family of Cell Density-Responsive Transcriptional Regulatorst. *J. Bacteriol.* **1994**, *176* (2), 269–275.

(5) Papenfort, K.; Bassler, B. L. Quorum Sensing Signal-Response Systems in Gram-Negative Bacteria. *Nat. Rev. Microbiol.* **2016**, *14* (9), 576–588.

(6) Homer, C. M.; Summers, D. K.; Goranov, A. I.; Clarke, S. C.; Wiesner, D. L.; Diedrich, J. K.; Moresco, J. J.; Toffaletti, D.; Upadhya, R.; Caradonna, I.; et al. Intracellular Action of a Secreted Peptide Required for Fungal Virulence. *Cell Host Microbe* **2016**, *19* (6), 849–864.

(7) Egland, P. G.; Palmer, R. J.; Kolenbrander, P. E. Interspecies Communication in *Streptococcus Gordonii-Veillonella Atypica* Biofilms: Signaling in Flow Conditions Requires Juxtaposition. *Proc. Natl. Acad. Sci.* **2004**, *101* (48), 16917–16922.

(8) McNab, R.; Lamont, R. J. Microbial Dinner-Party Conversations: The Role of LuxS in Interspecies Communication. *J. Med. Microbiol.* **2003**, *52* (7), 541–545.

(9) Xu, X.; He, Q.; Chen, C.; Zhang, C. Differential Communications between Fungi and Host Plants Revealed by Secretome Analysis of Phylogenetically Related Endophytic and Pathogenic Fungi. *PLoS One* **2016**, *11* (9), e0163368.

(10) Weber, H. Fatty Acid-Derived Signals in Plants. *Trends Plant Sci.* **2002**, *7* (5), 217–224.

(11) Porter, S. L.; Wadhams, G. H.; Armitage, J. P. Signal Processing in Complex Chemotaxis Pathways. *Nat. Rev. Microbiol.* **2011**, *9* (3), 153–165.

(12) Gudipaty, S. A.; Larsen, A. S.; Rensing, C.; Mcevoy, M. M. Regulation of Cu(I)/Ag(I) Efflux Genes in *Escherichia Coli* by the Sensor Kinase CusS. *FEMS Microbiol. Lett.* **2012**, *330* (1), 30–37.

(13) Potrykus, J.; Ballou, E. R.; Childers, D. S.; Brown, A. J. P. Conflicting Interests in the Pathogen-Host Tug of War: Fungal Micronutrient Scavenging Versus Mammalian Nutritional Immunity. *PLoS Pathog.* **2014**, *10* (3), 2–5.

(14) Merchant, S. S. The Elements of Plant Micronutrients. *Plant Physiol.* **2010**, *154* (2), 512–515.

(15) Merchant, S. S.; Allen, M. D.; Kropat, J.; Moseley, J. L.; Long, J. C.; Tottey, S.; Terauchi, A. M. Between a Rock and a Hard Place: Trace Element Nutrition in Chlamydomonas. *Biochim. Biophys. Acta* **2006**, *1763* (7), 578–594.

(16) Quinn, J. M. Two Copper-Responsive Elements Associated with the Chlamydomonas Cyc6 Gene Function as Targets for Transcriptional Activators. *Plant Cell Online* **1995**, *7* (5), 623–638.

(17) Kropat, J.; Tottey, S.; Birkenbihl, R. P.; Depege, N.; Huijser, P.; Merchant, S. A Regulator of Nutritional Copper Signaling in Chlamydomonas Is an SBP Domain Protein That Recognizes the GTAC Core of Copper Response Element. *Proc. Natl. Acad. Sci.* **2005**,

102 (51), 18730–18735.
(18) Sommer, F.; Kropat, J.; Malasarn, D.; Grossoehme, N. E.; Chen, X.; Giedroc, D. P.; Merchant, S. S. The CRR1 Nutritional Copper Sensor in *Chlamydomonas* Contains Two Distinct Metal-Responsive Domains. *Plant Cell* **2010**, *22* (12), 4098–4113.
(19) Quarmby, L. M.; Yueh, Y. G.; Cheshire, J. L.; Keller, L. R.; Snell, W. J.; Crain, R. C. Inositol Phospholipid Metabolism May Trigger Flagellar Excision in Chlamydomonas Reinhardtii. *J. Cell Biol.* **1992**, *116* (3), 737–744.
(20) Krouk, G.; Crawford, N. M.; Coruzzi, G. M.; Tsay, Y. Nitrate Signaling : Adaptation to Fluctuating Environments. *Curr. Opin. Plant Biol.* **2010**, *13* (3), 266–273.
(21) Duanmu, D.; Lagarias, J. C.; Merchant, S. S.; Yang, W.; Rockwell, N. C.; Martin, S. S.; Niyogi, K. K.; Dent, R. M.; Grossman, A. R.; Gallaher, S.; et al. Retrograde Bilin Signaling Enables Chlamydomonas Greening and Phototrophic Survival. *Proc. Natl. Acad. Sci.* **2013**, *110* (9), 3621–3626.
(22) de Montaigu, A.; Sanz-Luque, E.; Galván, A.; Fernández, E. A Soluble Guanylate Cyclase Mediates Negative Signaling by Ammonium on Expression of Nitrate Reductase in *Chlamydomonas*. *Plant Cell* **2010**, *22* (5), 1532–1548.
(23) Sanz-Luque, E.; Ocaña-Calahorro, F.; Llamas, A.; Galvan, A.; Fernandez, E. Nitric Oxide Controls Nitrate and Ammonium Assimilation in *Chlamydomonas Reinhardtii*. *J. Exp. Bot.* **2013**, *64* (11), 3373–3383.
(24) Sanz-Luque, E.; Ocaña-Calahorro, F.; Galván, A.; Fernández, E.; de Montaigu, A. Characterization of a Mutant Deficient for Ammonium and Nitric Oxide Signalling in the Model System *Chlamydomonas Reinhardtii*. *PLoS One* **2016**, *11* (5), e0155128.
(25) Kong, W. W.; Zhang, L. P.; Guo, K.; Liu, Z. P.; Yang, Z. M. Carbon Monoxide Improves Adaptation of Arabidopsis to Iron Deficiency. *Plant Biotechnol. J.* **2010**, *8* (1), 88–99.
(26) Liping, Z.; Hongbo, S.; Xiaohua, L.; Zhaopu, L. Gene Regulation of Iron-Deficiency Responses Is Associated with Carbon Monoxide and Heme Oxydase 1 in Chlamydomonas Reinhardtii. *PLoS One* **2013**, *8* (1), e0053835.
(27) Stone, J. R.; Marletta, M. A. Spectral and Kinetic Studies on the Activation of Soluble Guanylate Cyclase by Nitric Oxide. *Biochemistry* **1996**, *35* (4), 1093–1099.
(28) Zhao, Y.; Brandish, P.; Ballou, D.; Marletta, M. A. A Molecular Basis for Nitric Oxide Sensing by Soluble Guanylate Cyclase. *Proc. Natl. Acad. Sci. U. S. A.* **1999**, *96* (26), 14753–14758.
(29) Cary, S. P. L.; Winger, J. A.; Marletta, M. A. Tonic and Acute Nitric Oxide Signaling through Soluble Guanylate Cyclase Is Mediated by Nonheme Nitric Oxide, ATP, and GTP. *Proc. Natl. Acad. Sci.* **2005**, *102* (37), 13064–13069.
(30) Horst, B. G.; Marletta, M. A. Physiological Activation and Deactivation of Soluble Guanylate Cyclase. *Nitric Oxide* **2018**, *77*, 65–74.
(31) Derbyshire, E. R.; Marletta, M. A. Structure and Regulation of Soluble Guanylate Cyclase. *Annu. Rev. Biochem.* **2012**, *81* (1), 533–559.
(32) Buechler, W. A.; Nakane, M.; Murad, F. Expression of Soluble Guanylate Cyclase Activity Requires Both Enzyme Subunits. *Biochem. Biophys. Res. Commun.* **1991**, *174* (1), 351–357.
(33) Gray, J. M.; Karow, D. S.; Lu, H.; Chang, A. J.; Chang, J. S.; Ellis, R. E.; Marletta, M. A.; Bargmann, C. I. Oxygen Sensation and Social Feeding Mediated by a *C. Elegans* Guanylate Cyclase Homologue. *Nature* **2004**, *430* (6997), 317–322.
(34) Winter, M. B.; Woodward, J. J.; Marletta, M. A. *An Escherichia Coli Expression-Based*

Approach for Porphyrin Substitution in Heme Proteins; 2013; Vol. 987.
(35) Herzik, M. A.; Jonnalagadda, R.; Kuriyan, J.; Marletta, M. A. Structural Insights into the Role of Iron–Histidine Bond Cleavage in Nitric Oxide-Induced Activation of H-NOX Gas Sensor Proteins. *Proc. Natl. Acad. Sci.* **2014**, *111* (40), E4156–E4164.
(36) Goodstein, D. M.; Shu, S.; Howson, R.; Neupane, R.; Hayes, R. D.; Fazo, J.; Mitros, T.; Dirks, W.; Hellsten, U.; Putnam, N.; et al. Phytozome: A Comparative Platform for Green Plant Genomics. *Nucleic Acids Res.* **2012**, *40*, D1178–D1186.
(37) Urzica, E. I.; Casero, D.; Yamasaki, H.; Hsieh, S. I.; Adler, L. N.; Karpowicz, S. J.; Blaby-Haas, C. E.; Clarke, S. G.; Loo, J. A.; Pellegrini, M.; et al. Systems and *Trans*-System Level Analysis Identifies Conserved Iron Deficiency Responses in the Plant Lineage. *Plant Cell* **2012**, *24* (10), 3921–3948.
(38) Barr, I.; Guo, F. Pyridine Hemochromagen Assay for Determinig the Concentration of Heme Purified Protein Solutions. *Bio-Protocol* **2015**, *5* (18), e1594.
(39) Hespen, C. W.; Bruegger, J. J.; Phillips-Piro, C. M.; Marletta, M. A. Structural and Functional Evidence Indicates Selective Oxygen Signaling in Caldanaerobacter Subterraneus H-NOX. *ACS Chem. Biol.* **2016**, *11* (8), 2337–2346.
(40) Stone, J. R.; Marletta, M. A. Soluble Guanylate Cyclase from Bovine Lung: Activation with Nitric Oxide and Carbon Monoxide and Spectral Characterization of the Ferrous and Ferric States. *Biochemistry* **1994**, *33* (18), 5636–5640.
(41) Price, M. S.; Chao, L. Y.; Marletta, M. A. Shewanella Oneidensis MR-1 H-NOX Regulation of a Histidine Kinase By. *Biochemistry* **2007**, *46* (48), 13677–13683.
(42) Fernhoff, N. B.; Derbyshire, E. R.; Marletta, M. A. A Nitric Oxide/Cysteine Interaction Mediates the Activation of Soluble Guanylate Cyclase. *Proc. Natl. Acad. Sci.* **2009**, *106* (51), 21602–21607.
(43) Karl, D. M. Occurrence and Ecological Significance of GTP in the Ocean and in Microbial Cells. *Appl. Environ. Microbiol.* **1978**, *36* (2), 349–355.
(44) Traut, T. W. Physiological Concentrations of Purines and Pyrimidines. *Mol. Cell. Biochem.* **1994**, *140* (1), 1–22.
(45) Winger, J. A.; Derbyshire, E. R.; Marletta, M. A. Dissociation of Nitric Oxide from Soluble Guanylate Cyclase and Heme-Nitric Oxide/Oxygen Binding Domain Constructs. *J. Biol. Chem.* **2007**, *282* (2), 897–907.
(46) Derbyshire, E. R.; Gunn, A.; Ibrahim, M.; Spiro, T. G.; Britt, R. D.; Marletta, M. A. Characterization of Two Different Five-Coordinate Soluble Guanylate Cyclase. *Biochemistry* **2008**, *47* (12), 3892–3899.
(47) Gunn, A.; Derbyshire, E. R.; Marletta, M. A.; Britt, R. D. Conformationally Distinct Five-Coordinate Heme − NO Complexes of Soluble Guanylate Cyclase Elucidated by Multifrequency Electron Paramagnetic Resonance (EPR). *Biochemistry* **2012**, *51* (42), 8384–8390.
(48) Guo, Y.; Suess, D. L. M.; Herzik, M. A.; Iavarone, A. T.; Britt, R. D.; Marletta, M. A. Regulation of Nitric Oxide Signaling by Formation of a Distal Receptor–Ligand Complex. *Nat. Chem. Biol.* **2017**, *13* (12), 1216–1221.
(49) Hespen, C. W.; Bruegger, J. J.; Guo, Y.; Marletta, M. A. Native Alanine Substitution in the Glycine Hinge Modulates Conformational Flexibility of Heme Nitric Oxide/Oxygen (H-NOX) Sensing Proteins. *ACS Chem. Biol.* **2018**, *13*, 1631–1639.
(50) Martin, E.; Berka, V.; Bogatenkova, E.; Murad, F.; Tsai, A. L. Ligand Selectivity of Soluble Guanylyl Cyclase: Effect of the Hydrogen-Bonding Tyrosine in the Distal Heme

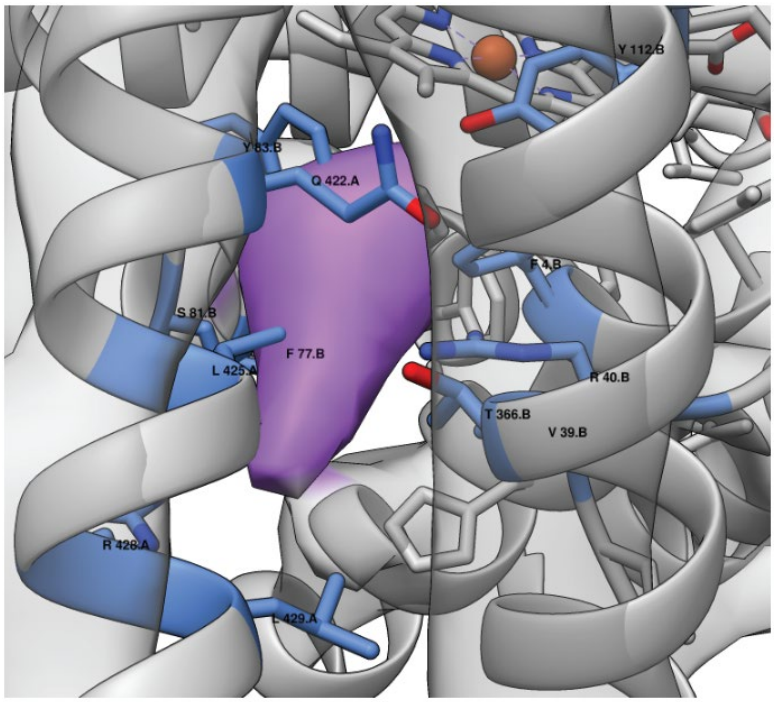
Pocket on Binding of Oxygen, Nitric Oxide, and Carbon Monoxide. *J. Biol. Chem.* **2006**, *281* (38), 27836–27845.
(51) He, Y.; Tang, R.-H.; Hao, Y.; Stevens, R. D.; Cook, C. W.; Ahn, S. M.; Jing, L.; Yang, Z.; Chen, L.; Guo, F.; et al. Nitric Oxide Represses the *Arabidopsis* Floral Transition. *Science (80-.).* **2004**, *305* (5692), 1968–1971.
(52) Salmi, M. L.; Morris, K. E.; Roux, S. J.; Porterfield, D. M. Nitric Oxide and CGMP Signaling in Calcium-Dependent Development of Cell Polarity in *Ceratopteris Richardii*. *Plant Physiol.* **2007**, *144* (1), 94–104.
(53) Wei, L.; Derrien, B.; Gautier, A.; Houille-Vernes, L.; Boulouis, A.; Saint-Marcoux, D.; Malnoë, A.; Rappaport, F.; de Vitry, C.; Vallon, O.; et al. Nitric Oxide-Triggered Remodeling of Chloroplast Bioenergetics and Thylakoid Proteins upon Nitrogen Starvation in *Chlamydomonas Reinhardtii*. *Plant Cell* **2014**, *26* (1), 353–372.
(54) Düner, M.; Lambertz, J.; Mügge, C.; Hemschemeier, A. The Soluble Guanylate Cyclase CYG12 Is Required for the Acclimation to Hypoxia and Trophic Regimes in *Chlamydomonas Reinhardtii*. *Plant J.* **2018**, *93* (2), 311–337.

CHAPTER 5

EXPERIMENTS PERTAINING TO CHAPTER 4, SUMMARY, AND FUTURE DIRECTIONS

Summary

The work presented in this thesis has answered outstanding questions concerning the biochemistry of gas receptor proteins and in the process discovered a new class of these receptors. In Chapter 2, the full-length structure of *Manduca sexta* sGC was described and the binding pocket of small molecule stimulators was found to be directly between the β H-NOX domain and the two CC domains. Chapter 3 presented studies that sought to uncover the role of the second NO binding site, which will lead to the outlining of new and potentially impactful experiments needed to answer this question. Chapter 4 described a new class of soluble guanylate cyclase proteins that are most activated by carbon monoxide. The remainder of this thesis will be devoted to experiments performed that were not published previously relating to work done in Chapter 4, as well as a treatise concerning future directions of the biochemistry of gas receptor proteins.



Introduction

The unicellular alga *Chlamydomonas reinhardtii* has been shown to possess six genes that encode for proteins with the same domain structure as the mammalian nitric oxide receptor, soluble guanylate cyclase (sGC).[1] Chapter 4 of this thesis described the biochemistry of one these gene products, Cyg11, and showed that this homologue of sGC is more active when carbon monoxide (CO) is bound than when nitric oxide (NO) is bound.[2] This is the first sGC to be characterized with this activation profile, which brings credence to the idea of CO functioning as a signaling molecule in a eukaryotic organism. However, there are outstanding questions still to be addressed both within this story and within the larger study of gas receptor proteins.

With six total genes that resemble the mammalian sGCs architecture in the genome of *C. reinhardtii*, yet sequence analysis unable to sort these genes into α-like and β-like genes, there exists 21 possibilities of homodimeric and heterodimeric oligomers. Some heterodimeric combinations seem more likely than others, such as Cyg11-Cyg12, which exist next to each other in the genome. Different combinations of these genes were co-expressed to determine the ligand-stimulated activity of these different oligomers.

Despite Cyg11 from *C. reinhardtii* possessing the lowest $K_{D,CO}$ of any previous biochemically characterized sGC at 2.5 μM^2, the $K_{D,NO}$ is presumed to be many orders of magnitude lower, as is common for ferrous heme proteins.[3] This fact is often cited as the reason CO is not a good candidate for a signaling gas, and why organisms that possess both NO-specific and O_2-specific sGCs transcriptionally regulate their expression so they are never expressed together in the same tissues[4] The evidence presented in Chapter 4 showing that Cyg11 potentially is activated by CO *in vivo* raises the question about how large the the $K_{D,NO}$ for Cyg11 is.

Additionally, the datasets that showed *cyg11* to be increasing in transcription as the iron level decreased in the culturing media were taken from large transcriptomics datasets and are worth repeating. This chapter outlines the experiments conducted with regards to heterodimers, the NO on-rate of Cyg11, and the levels of mRNA present for all six *cyg* genes.

Materials and Methods

Materials

All chemicals were purchased from commercial vendors and used without further purification unless otherwise noted. Primers were obtained from Integrated DNA Technologies (Coralville, IA). Gibson Master Mix was purchased from New England Biolabs (Ipswich, MA). PrimeSTAR Max DNA polymerase and His60 Ni Superflow resin was purchased from Takara Bio USA (Mountain View, CA). DNA purification kits were purchased from Qiagen (Germantown, MD). Terrific Broth media powder, ampicillin, and 4-(2-aminoethyl)benzenesulfonyl fluoride hydrochloride (AEBSF) were purchased from Research Products International (Mount Prospect, IL). Hemin chloride, sodium dithionite ($Na_2S_2O_4$), DNAse I from bovine pancreas, β-mercaptoethanol, H_2NaPO_4, $NaCH_3CO_2$, $Mg(CH_3CO_2)_2$, and guanosine 5′-triphosphate sodium salt (GTP) (>95%, HPLC) were purchased from Millipore Sigma (Burlington, MA). Isopropyl β-D-1-thiogalactopyranoside (IPTG), 4-(2-Hydroxyethyl)piperazine-1-ethanesulfonic acid

(HEPES), and Pierce BCA Protein Assay Kit, were purchased from Thermo Fisher Scientific (Waltham, MA). Benzamidine, lysozyme from chicken egg white, and sodium chloride (NaCl) were purchased from VWR Life Science (Visalia, CA). Dithiothreitol (DTT) was purchased from Bachem (Bubendorf, Switzerland). Imidazole was purchased from Oakwood Chemical (West Columbia, SC). Vivaspin spin concentrators were purchased from Sartorius (Concord, CA). PD-10 desalting columns were purchased from GE Healthcare (Pittsburgh, PA). DEA NONOate was purchased from Cayman Chemical Company (Ann Arbor, MI). Nitric oxide (NO, >99.5%) gas and carbon monoxide (CO, >99%) gas were purchased from Praxair Inc. (Danbury, CT). NO gas was bubbled through 10 M NaOH to remove higher oxides prior to use.

Construction of Expression Plasmids

The Core cDNA library of *Chlamydomonas reinhardtii* was purchased from the Chlamydomonas Resource Center (University of Minnesota). Primers were designed against the 3′ and 5′ UTR of *cyg11* (transcript name: Cre07.g320700.t1.1) Fwd: 5′-GGATCCATCGATGCTTAGGAGGTC GCCATGCAGTTCATTCTTAGT -3′, Rev: 5′-CTGCTCATGTTTGACAGCTTATCATCG CGTGTCCGCGAGCAATGGTAA -3′. The underlined regions of the primers anneal to 3′ and 5′ UTR of Cyg11. The other *C. reinhardtii* genes were cloned out of the *C. reinhardtii* library by Jonathan Winger, PhD. The transcript names for *cyg12*, *cyg15*, and *cyg38* are Cre07.g320750.t1.1, Cre07.g321150.t1.1, Cre16.g649050.t2.1, respectively. The purified PCR product was subcloned into a linearized pCW vector with Gibson Master Mix, and the product was verified by DNA sequencing (UC Berkeley DNA Sequencing Facility). The *C. reinhardtii* sGC predicted protein sequence were aligned to the β-subunit protein sequence of *Rattus norvegicus* (*Rn*) sGC (Uniprot: Q02153-1) using Clustal Omega (EMBL-EBI). The crystal structure of the *Homo sapiens* catalytic domain (99% identical to *Rn*, PDB ID: 4NI2) was used to identify the last ordered residue, and thus the N-terminal amino acids were subcloned from the full-length *C. reinhardtii* sGC genes. The final homodimer expression plasmids include either a hexahistidine tag or a strep tag at the C-terminal of the protein, as the N-terminus of H-NOX domains are buried in the heme binding pocket and cannot be extended. The six co-expression plasmids were made by cloning the genes out of the subcloned plasmids and into a modified pET_duet plasmid. The final expression and co-expression constructs listed from 5′ to 3′ are pCW_Cyg11$_{615}$_His$_6$, pCW_Cyg12$_{651}$_His6, pCW_Cyg15$_{600}$_His$_6$, pCW_Cyg38$_{583}$_His$_6$, pET_Cyg11$_{615}$_His$_6$_Cyg12$_{651}$_ST, pET_Cyg11$_{615}$_His$_6$_Cyg15$_{600}$_ST, pET_Cyg11$_{615}$_His$_6$_Cyg38$_{583}$_ST, pET_Cyg15$_{600}$_His$_6$_Cyg12$_{651}$_ST, pET_Cyg38$_{583}$_His$_6$_Cyg12$_{651}$_ST, and pET_Cyg15$_{600}$_His$_6$_Cyg38$_{583}$_ST. The full peptide sequences of constructs used in this chapter is included in Figure C.1.

Protein Expression and Purification

Please see the Chapter 4 Materials and Methods section for the protein expression and purification of the four homodimeric constructs. The following changes were made for heterodimeric constructs. Expression, cell lysis, clarification, and a preliminary purification by IMAC with His$_{60}$ was performed as before. The elution was buffer exchanged to remove the imidazole, and then applied to a column with Streptactin resin (Qiagen) and eluted off with increasing concentrations of desthiobiotin up to 2.5 mM (Millipore Sigma). Western blots were probed for both His tagged and strep tagged proteins by applying mouse anti-His$_6$ Ab followed by a goat anti-mouse HRP conjugate Ab (Life Technologies) as before. The blots were stripped by the addition of two 5 mL washes for five minutes each with Stripping Buffer (200mM Glycine, 1% Tween20, 0.1% SDS).

The blot was re-blocked followed by the application of a strep tag antibody conjugated to HRP (BioRad) at a 1:5000 dilution and detected with Classico substrate (Millipore Sigma).

Activity Assays and Quantification

Steady-state kinetics for *Cr* sGC were measured by quantifying the amount of cGMP produced in duplicate endpoint activity assays, performed in at least biological triplicate. Samples were reduced in an anaerobic glove bag with 5 mM $Na_2S_2O_4$ for 15 minutes, and then desalted with a Minitrap column equilibrated with Buffer E (50 mM Hepes, pH 7.5, 150 mM NaCl, 5% (v/v) glycerol, degassed. The protein concentration was determined using the protein A_{280} peak with extinction coefficients generated from ExPASy with all cysteines reduced. CO (g) saturated buffer (950 μM) was prepared by sparging 3 mL of Buffer E in a Reacti-vial for 15 mL with CO (g). Gas-bound complexes of the samples were formed by adding DEA NONOate to a final concentration of 500 μM or CO (g) saturated buffer to a concentration of 425 μM. Activity assays were conducted at 25 °C and pH 7.5 in Buffer E, supplemented with 5 mM DTT and 5 mM $MgCl_2$. Reactions were initiated with 2 mM GTP and timepoints from were quenched with 125 mM $Zn(CH_3CO_2)_2$, followed by 125 mM $Na_2(CO_3)$ to adjust to pH 7. Samples were frozen at –80 °C until quantification. Quenched assays were thawed, and the zinc precipitate was spun down for 10 min at 23,150 *g*. The reactions were diluted by one to three orders of magnitude, and the cGMP was quantified in duplicate using an extracellular cGMP Enzyme Linked Immunosorbent Assay, following the manufacturer's instructions (Enzo Life Sciences). Concentrations of cGMP were determined from a standard curve, generated over 0.16–500 pmol/mL. Initial rates were calculated from the linear phase of the time course, where 5–10% of the GTP substrate was consumed.

Gas Binding Kinetics of NO

NO association rates (k_{on}) were determined for ferrous Cyg11. Samples (approximately 2 μM) were transferred in gas-tight syringes into the stopped-flow system that was made anaerobic by incubating the sample-handling unit with 10 mM $Na_2S_2O_4$ for 30 minutes. Dilutions of NO-saturated buffer were made on the sample handling unit in gas-tight syringes by diluting NO-saturated buffer with anaerobic Buffer H (50 mM HEPES, pH 7.5, 150 mM NaCl, 0.22 μm filtered) via a three-way valve. Reactions were initiated by mixing equal volumes of protein solution and NO-saturated buffer at 5 °C for a final protein concentration of 1 μM.

Transient Absorption Spectroscopy

The Cyg11 homodimer was reduced with 50 mM $Na_2S_2O_4$ for 15 minutes at room temperature. The reduced sample was desalted over a Miditrap G-25 column equilibrated in Buffer I (50 mM K_2HPO_4 pH 8, 200 mM NaCl, 5% glycerol, 0.22 μM filtered, degassed), adjusted to 20 μM, and snap frozen in liquid nitrogen. Samples were transported to either Caltech or Lawrence Berkeley National Laboratory on dry ice. Buffer I was degassed by pulling vacuum and backfilling with Ar (g) 10 times. NO (g) saturated buffer (1.9 mM) was prepared by sparging 3 mL of anaerobic Buffer I for 15 minutes in a Reacti-vial. In a glove box, four different samples of Cyg11 with different NO (g) concentrations (200 μM, 100 μM, 50 μM, and 25 μM) were prepared. UV-Vis absorption spectra were taken prior to and after transient absorption spectroscopy.

Transient absorption spectroscopy was performed with a Nd:YAG laser tuned to the second harmonic in order to provide more power (~100 mJ) over timescales out to: 2 μs, 100 μs, 10 ms, and 500 ms. Two different probe lights were used for different time scales, a xenon arc lamp was

used for timescale shorter than 10 ms and a tungsten lamp was used for timescales longer than 500 ms.

Strains, Media, and Growth Conditions

Chlamydomonas reinhardtii WT(+) was received as a gift from the Niyogi lab (UC Berkeley). The following stock solutions were prepared and autoclaved, with the stock concentration indicated in parenthesis: FBS (80x): 600 mM NH_4Cl, 27 mM $CaCl_2$, 32 mM $MgSO_4$; KPO_4 (200x): 200 mM K_2HPO_4; Modified Trace Metals solution (200x): 27 mM EDTA, bring to pH 6.5, 3.6 mM $FeSO_4$, 15 mM $ZnSO_4$, 37 mM H_3BO_3, 5.2 mM $MnCl_2$, 1.3 mM $CuCl_2$, 1.1 mM H_2MoO_4, and 1.3 mM $CoCl_2$; HSBeji (200x): 1,900 mM NH_4Cl, 14 mM $CaCl_2$, 16 mM $MgSO_4$; HSPhos (200x): 2,710 mM K_2HPO_4. TAP media was prepared by preparing 20 mM Tris base in 975 mL of Milli Q dH_2O followed by the addition of 1 mL of glacial acetic acid for a final concentration of 17.5 mM of acetate. Next, 12.5 mL of FBS (80x) and 5 mL each of KPO_4 (200x) and Modified Trace Metals solution (200x) were added under sterile conditions. To prepare TAP-agar, 1.5% (w/v) of Difco agar was added at this stage. The TAP media or TAP-agar was autoclaved for 45 minutes at 121 °C. HS media was prepared by adding 5 mL each of HSBeji (200x), HSPhos (200x), and Modified Trace Metals solution (200x) to 985 mL of Milli Q dH_2O, and autoclaving for 45 minutes at 121 °C.

C. reinhardtii was maintained on TAP-agar plates grown with 20-100 μmol m^{-2} s^{-1} of photons of light and plates were restreaked every 3-4 days. It should be noted that this is considered dim light for *C. reinhardtii*, as typically algae are grown in high light (1,200 μmol m^{-2} s^{-1} of photons).[5] Cells were transferred from plates to suspension culture by scraping a pea-size amount of cells into 50 mL of the appropriate media and shaking for 48 hours at room temperature for the culture to acclimate. Cultures were synchronized with 12 hours in the dark (wrapped in foil), followed by 12 hours in the light. Three different growth conditions were used: mixotrophic (TAP + light), photosynthetic (HS media + light), and dark (TAP + dark). *C. reinhardtii* cells were counted with a Coulter Counter by diluting 50 μL of culture into 10 mL of isoton solution, and adjusting the OD750 to the Coulter Counter readings.

RNA extraction, cDNA preparation, and RT-qPCR quantification

All steps were performed at 4 °C unless otherwise noted. RNA was isolated from mid-log phase *C reinhardtii* cultures. 2 x 10^7 cells were lysed with 1 mL of Trizol reagent (Thermo Fisher), then boiled at 95 °C for 5 min. Debris was pelleted by centrifuging the sample at 12,000 *g*, 4 °C, for 10 minutes. Nucleic acids were separated with the addition of 0.2 mL chloroform, the sample was vortexed well and then centrifuged at 12,000 *g*, 4 °C, for 15 minutes. A 500 μL fraction of the aqueous layer (top) was added to 0.5 mL ice cold isopropanol, incubated at RT for 10 minutes, and then centrifuged at 12,000 *g*, 4 °C, for 10 minutes. The resulting pellet was washed with 1 mL of 75% ethanol and dissolved in 40 μL RNAse-free dH_2O. Any remaining DNA was digested with a DNase kit following manufacturer's instructions (Ambion). Prepare cDNA with cDNA synthesis kit following manufacturer's instructions (Invitrogen SuperScript II). Primer efficiency was checked by diluted either plasmid or *C. reinhardtii* gDNA from 3000 pg down to 0.3 pg per well. cDNA was quantified with Maxima SYBR Green/Fluorescein qPCR Master Mix following the manufacturer's instructions (Thermo Fisher). Controls included no nucleic acid with all primer pairs and RNA without cDNA synthesis for all primer pairs.

Results and Discussion

Meta-Analysis of Cr sGC Transcriptomics

While Chapter 4 showed that homodimeric Cyg11 is active and stimulated by CO, there is the possibility that other heterodimeric sGCs could form in *Chlamydomons reinhardtii*. These heterodimeric species could have different activation profiles than Cyg11. It was hypothesized that if heterodimers formed in *C. reinhardtii*, there should be equivalent upregulation of these two genes in response to a stimulus. Thus, a meta-analysis of RNAseq datasets for gene transcription was performed using data that has been compiled by JGI Phytozome (Figure 5.1A). Figure 5.1B shows a selection of experiments that result in the upregulation of some of the six *Cr* sGCs, with the left panel comparing light and dark growth conditions and the right panel comparing mixotrophic (acetate + light) and photosynthetic (no acetate + light) conditions. While *cyg12* was upregulated in both dark and photosynthetic conditions and *cyg38* was upregulated in photosynthetic conditions only, the only two sGCs that seemed to be upregulated concomitantly were *cyg11* and *cyg12* in photosynthetic conditions. Additionally, the genes for *cyg11* and *cyg12* are found in direct proximity to each other on chromosome 7, lending additional support to the hypothesis that the product of these two genes could form a heterodimeric species (Figure 5.1C). Therefore, a coexpression vector was designed in order to express both Cyg11 and Cyg12 from the same plasmid to test the ligand-stimulated activity for this heterodimer.

Activity Assays of Cr sGC Heterodimers

The co-expression construct contained $Cyg11_{615}$ with a His_6 tag and $Cyg12_{651}$ with a Strep tag. This design allowed for the separation of any Cyg11 or Cyg12 homodimers from the heterodimer by passing lysate first through a Ni^{2+} IMAC column followed by a streptactin column. Figure C.2 shows SDS-PAGE purification gels and Western blots with both anti-His and anti-Strep antibodies. The presence of both tags in the final elution volume is indicative of the successful purification of the heterodimer Cyg11-Cyg12. Next, the ligand-stimulated activity of the Cyg11-Cyg12 heterodimer was interrogated with *in vitro* cGMP assays, followed by cGMP quantification with a cGMP-specific ELISA. Figure 5.2 shows the amount of cGMP produced by the heterodimeric construct both after the Ni^{2+} IMAC column and the Strep column compared to a *Rn* sGC WT enzyme control. There is no discernable ligand-stimulated activity for this heterodimeric construct. Thus, the other five heterodimeric constructs were cloned but were not expressed to determine if they showed the same result.

NO On Rate for Cyg11

Cyg11 has the lowest $K_{D,CO}$ for any biochemically characterized sGC, suggesting that CO is possibly the physiological ligand. However, NO typically has a very low $K_{D,NO}$ for ferrous heme cofactors. For example, the $K_{D,NO}$ for *Rn* sGC is estimated to be $< 1.2 \times 10^{-12}$ M.[3] Chapter 4 reported the off-rate for NO from Cyg11 was two orders of magnitude faster than the off-rate for NO from *Rn* sGC, however in order to compare dissociation constants both the off-rate and the on-rate are required. Similar to what has been done previously, stopped-flow spectroscopy was used with varying NO concentrations in attempts to obtain the $k_{on,NO}$. We would expect to see the Soret peak undergo two transitions.[3] The first transition begins with the unliganded peak centered at 431 nm that converts to a 6-coordinate ferrous nitrosyl peak at 420 nm with the proximal histidine still bound. The second transition begins with the 6-coordinate ferrous nitrosyl and

converts to a 5-coordinate ferrous nitrosyl peak at 399 nm with the cleavage of the iron histidine bond. Figure 5.3A shows the change of the Soret peak at three different

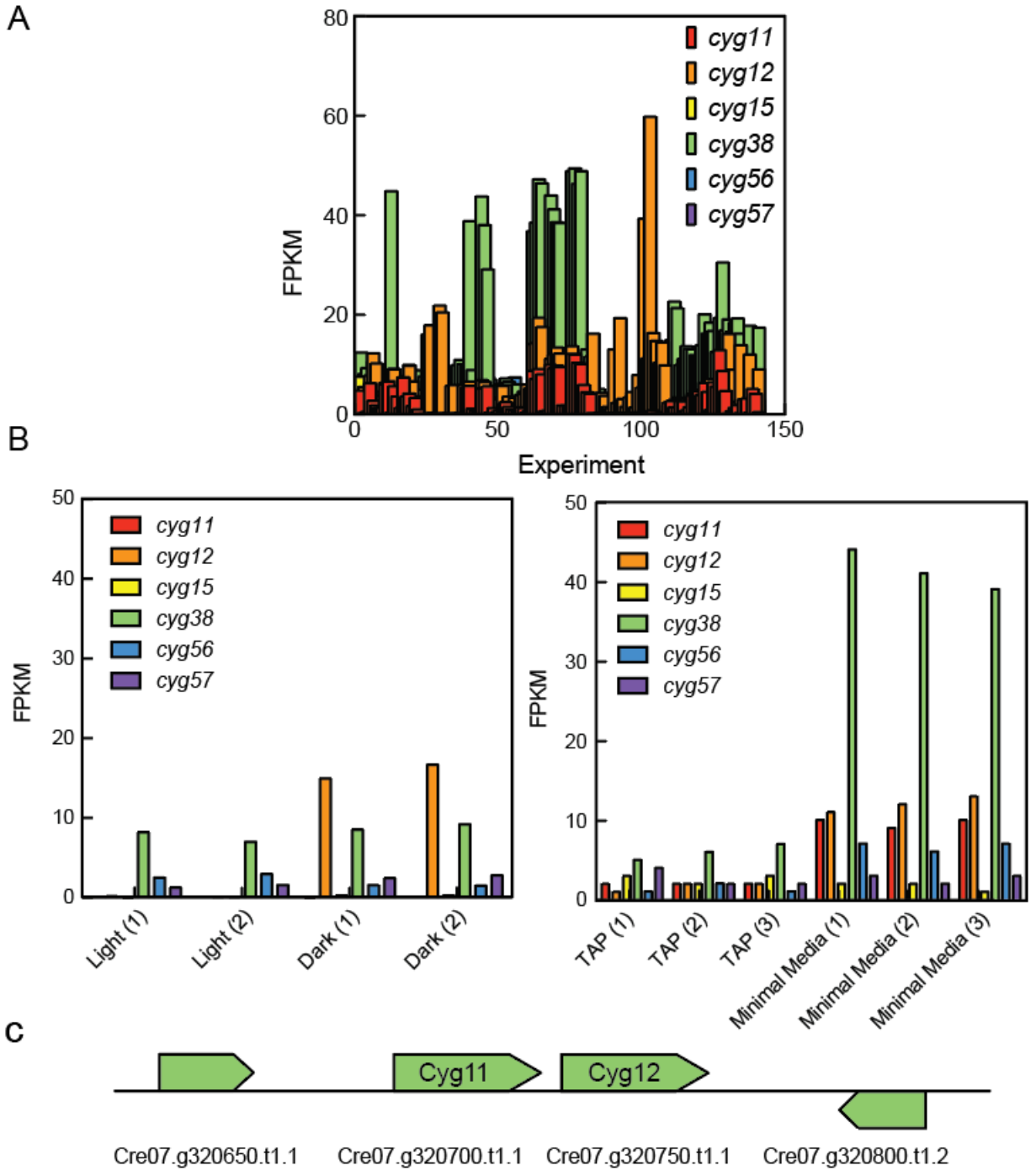


Figure 5.1 – Transcriptomic analysis of *C. reinhardtii* sGCs. Data in this figure is from JGI Phytozome (https://phytozome.jgi.doe.gov/pz/portal.html) A) mRNA transcript level in FPKMs of the six *Cr* sGCs in over 130 different experiments. B) Selected experiments that show differential expression in one or more *Cr* sGCs. C) Schematic of *Cr* chromosome 7 where Cyg11 and Cyg12 are directly adjacent to each other in the genome. The transcript number of the genome neighborhood is listed below the genes.

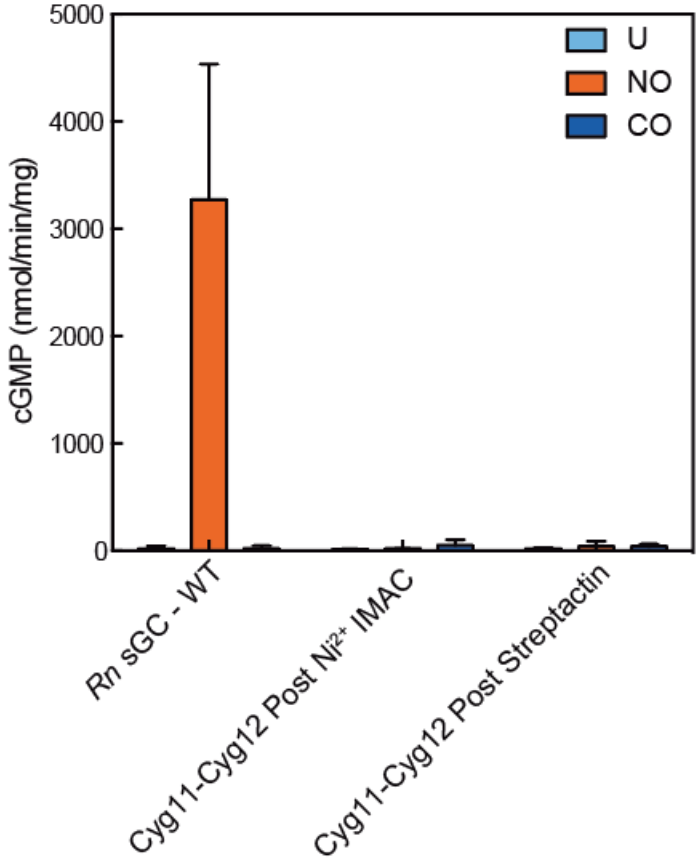


Figure 5.2 – Endpoint activity assay of *Rn* sGC and Cyg11-Cyg12 co-expression. No significant activity was observed for the heterodimeric construct of Cyg11-Cyg12. Other heterodimeric constructs were cloned but never expressed and purified.

wavelengths over time. Only the transition from 420 nm to 399 nm is observed, indicating that the association between NO and Cyg11 is occurring within the dead time of the instrument. Even when the NO concentration is decreased to 29 μM, only one transition is observed (Figure 5.3B). These transitions at 431 nm can be fit by double exponential decay functions and Figure 5.3C indicates that only one of the parameters is concentration dependent. However, without observing the transition from 431nm to 420 nm, only a lower limit for $k_{on,NO}$ of 7.43 x 10^4 M^{-1} s^{-1} can be set based on the faster transition in Figure 5.3C (k_{fast}). Combining this with the $k_{off,NO}$ measured in Chapter 4, an upper limit for $K_{D,NO}$ for Cyg11 is 1.9 μM, which is still lower than the $K_{D,CO}$ for Cyg11 (2.5 μM).

To obtain a NO on-rate for Cyg11 at timescales shorter than milliseconds, we turned to transient absorption spectroscopy. Ultra-fast dynamics of NO binding to the *Rn* sGC heme have been reported, albeit on a laser setup made in-house.[6] Two different Nd-YAG flash photolysis laser setups were used, one at Caltech (Gray lab) and one at Lawrence Berkeley National Lab (Frei lab). Experiments began with a sample of Cyg11 with NO bound to the heme. A 532 nm pulse

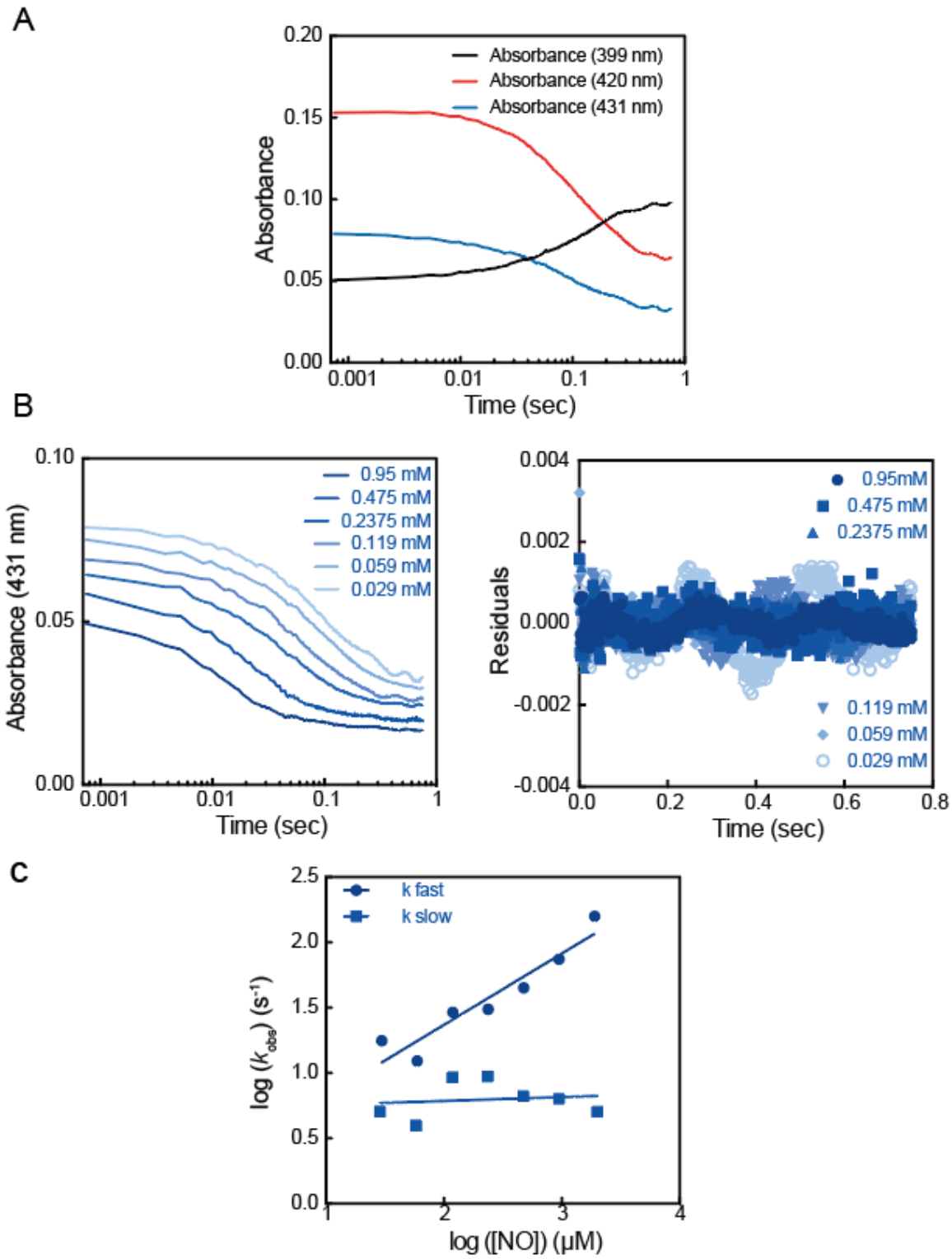


Figure 5.3 – Cyg11 NO on-rates. A) The change in three different wavelengths is plotted over time as unliganded Cyg11 homodimer is mixed with NO (0.029 mM). B) The change in absorbance at 431 nm corresponding to the maximum peak of the unliganded Soret band is plotted over time and fit to double exponential decay functions, the residuals of which are shown on the right. C) The fast biexponential rate constant varied over different concentrations of NO while the slow biexponential rate constant was constant over a range of NO concentrations.

would rupture the bond between the iron and NO and changes in the Soret band would be observed on timescales long enough that we were confident that a bimolecular process was occurring (and the NO was not remaining within the protein to rebound back onto the heme). While some quenches and absorptions were observed, the many different processes occurring at 431 nm and the proportion of protein that was predicted to undergo a bimolecular process (~5%) in addition to indications of sample degradation prevented interpretable data from being obtained.

Targeted Transcriptomics of Cr sGCs

In order to examine the relevance of gas signaling in cultures of *C. reinhardtii*, attempts were made to examine the properties and activation of natively expressed sGCs. *C. reinhardtii* 4A mt(+) was grown in different conditions to interrogate if any condition increased the expression of *Cr* sGCs. As noted above, *C. reinhardtii* can either use acetate as a carbon source or use photosynthesis to grow on CO_2, so three different growth conditions were used: mixotrophic (acetate + light), photosynthetic (no acetate + light), and dark (acetate + dark). A growth curve for *C. reinhardtii* in these three different conditions is shown in Figure 5.4. For the mixotrophic and photosynthetic conditions, cultures were in fact grown under dim light (20-100 μmol m^{-2} s^{-1}).

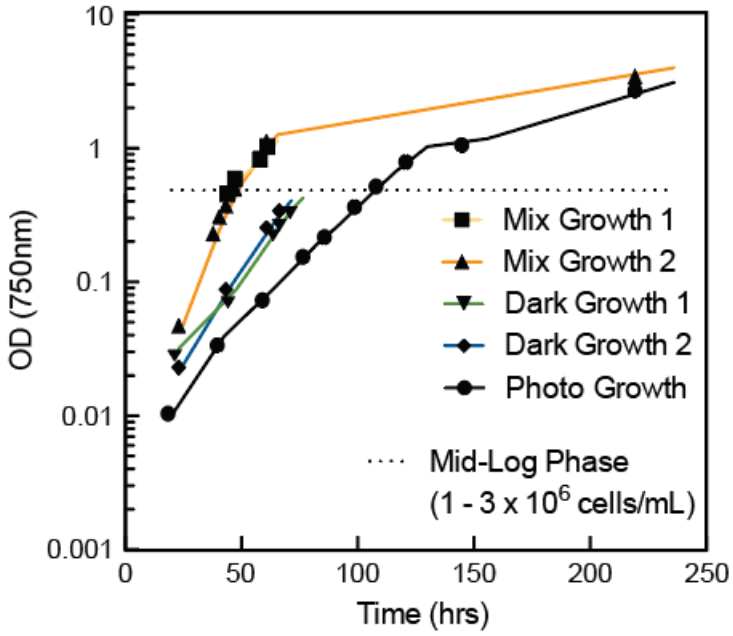


Figure 5.4 – *C. reinhardtii* growth curves measured by optical density. Growth curves of *C. reinhardtii* are plotted in from cultures grown in different growth conditions. The optical density at 750 nm was calibrated to cell counts measured by a Coulter Counter.

Initial investigations focused on mRNA levels for sGCs in different conditions. Two different reference genes were investigated for stable expression over all three conditions, *CβLP* (*rack1*, receptor of activated protein kinase C) and ubiquitin (*UBQ2*).[7,8] Figure 5.5A shows the relative *CβLP* transcription remained constant in the three different growth conditions, and thus *CβLP* was used to normalize all transcription in future experiments. Only primer pairs for three of the six *Cr* sGCs (*cyg12*, *cyg38*, and *cyg56*) were shown to both be both specific and efficient, and the relative transcription of these genes is shown in Figures 5.5B-D. Transcription of both *cyg12* and *cyg56* showed significant increases in both the dark and photosynthetic growth conditions,

suggesting that gas sensing is potentially more important in times when the preferred carbon source is not readily available (keeping in mind that cultures were not given sufficient light in the

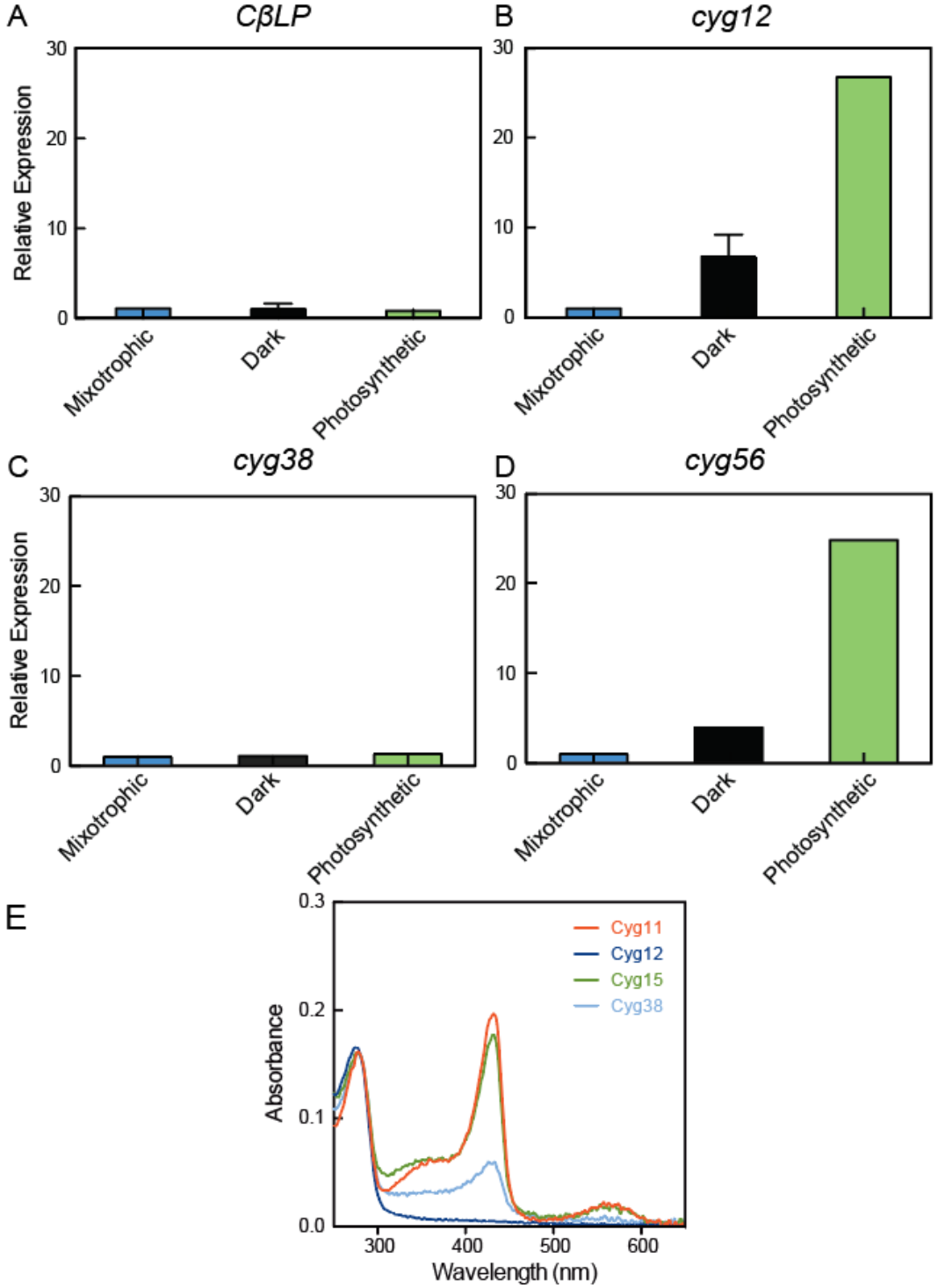


Figure 5.5 – *Cr* sGC transcriptomics and heme incorporation. A-D) Relative expression of one reference gene *CβLP* and three *Cr* sGCs was quantified by RT-qPCR and the data are shown. Both *cyg12* and *cyg56* are upregulated in photosynthetic growth conditions. Increases in *cyg12* transcription were expected based on data collected from JGI Phytozome. Increases in *cyg56* transcription were not expected. E) Relative heme incorporation for homodimers of Cyg11, Cyg12, Cyg15, and Cyg38. Chapter 4 discussed the heme incorporation for Cyg11 to be one heme per homodimer.

photosynthetic growth condition). The biochemical data collected for the homodimers of Cyg12 and Cyg38 revealed that neither homodimer had high levels of heme incorporation (Figure 5.5E), and no biochemical data had been collected for Cyg56. The fact that transcriptomics of *cyg11* was not able to be collected leaves gaps between the biochemical data and the transcriptomic data that potentially could be complementary in the future.

Conclusion

Chlamydomonas reinhardtii possesses six genes whose gene products are homologous to the mammalian nitric oxide receptor, sGC. One of those gene produces, Cyg11, has been found to be the first sGC homologue that is activated more by CO than by NO. However, there are lingering questions that work in this chapter sought to address. Despite co-expression analysis that did not predict any of these *Cr* sGCs to be co-expressed, the potential for homodimerization of these genes *in vivo* still exists. Targeted transcriptomic analysis of the complement of these genes in different growth conditions is still warranted to understand the transcription patterns and gain insight into the signaling pathways the gene products are a part of. Cyg11 must have a lower $K_{D,NO}$ than $K_{D,CO}$, however whether or not the alga produces and/or encounters both of these gases at the same time and at what concentration is not known. The results of the experiments presented here do not constitute a completed story, and further work is necessary to gain a complete understanding on the biochemical gaseous signaling pathways in *Chlamydomonas reinhardtii*.

Cumulative Future Directions

The work presented in this thesis describes the full-length structure of the eukaryotic nitric oxide receptor soluble guanylate cyclase, the investigation into cysteine residues as the potential second site of nitric oxide binding on sGC, and the discovery of a new class of soluble guanylate cyclases that are more activated by CO than by NO. Many new areas of research can now be initiated based off of results presented herein.

The full-length structure of sGC presented in this thesis, in combination with a high-resolution structure of the human sGC that was recently published[9], allow for the first time targeted mutational analysis of sGC to understand the activation mechanism. All residue numbers in this section will refer to the *Hs* sequence from PDB IDs 6JT0 and 6JT2 unless explicitly stated. Analyses of the structures begins with a close examination of the β H-NOX domain in the inactive and active states. The distance between the ε2 nitrogen of βH105 and the iron of the heme cofactor in the inactive state is 3.6 Å, however that distance only increases 0.3 Å to 3.9 Å in the activated structure (Figure 5.6). This is in contrast to bacterial H-NOX domains in different gas-bound states. For example, the H-NOX protein from *Shewanella oneidensis* has a distance between the ε2 nitrogen of H103 and the iron of 2.2 Å in the unliganded state. However that distance increases to 10.4 Å in the NO-bound state.[10] This implies that there is not the same rotation of the αF helix in sGCs that there is in H-NOX domains. However, in both *Hs* sGC and *So* H-NOX structures, the N-terminal subdomain of the H-NOX rotates ~4° away from the C-terminal subdomain. It is

possible this motion is the critical conserved motion of H-NOX domains, as O_2 sensing H-NOX domains also show this relative motion between N- and C-terminal subdomains.[11]

The β H-NOX in the inactive state contains many residues that make polar contacts with residues of other domains (Figure 5.7). Specifically, βD106 contacts βR258, both of which are strictly conserved (signified by the underline) from the SSN analysis discussed in Chapter 3, and βD106 has been shown to be critical for sGC activation.[9,12] Additionally, βR258 is nestled within a network of many polar contacts, interacting with βR305 (R/K strictly conserved), βK307, βE370, and βD374, with more polar residues existing in a second sphere that supports this inner sphere surrounding βR258 (βY363, αD411, and αQ418).

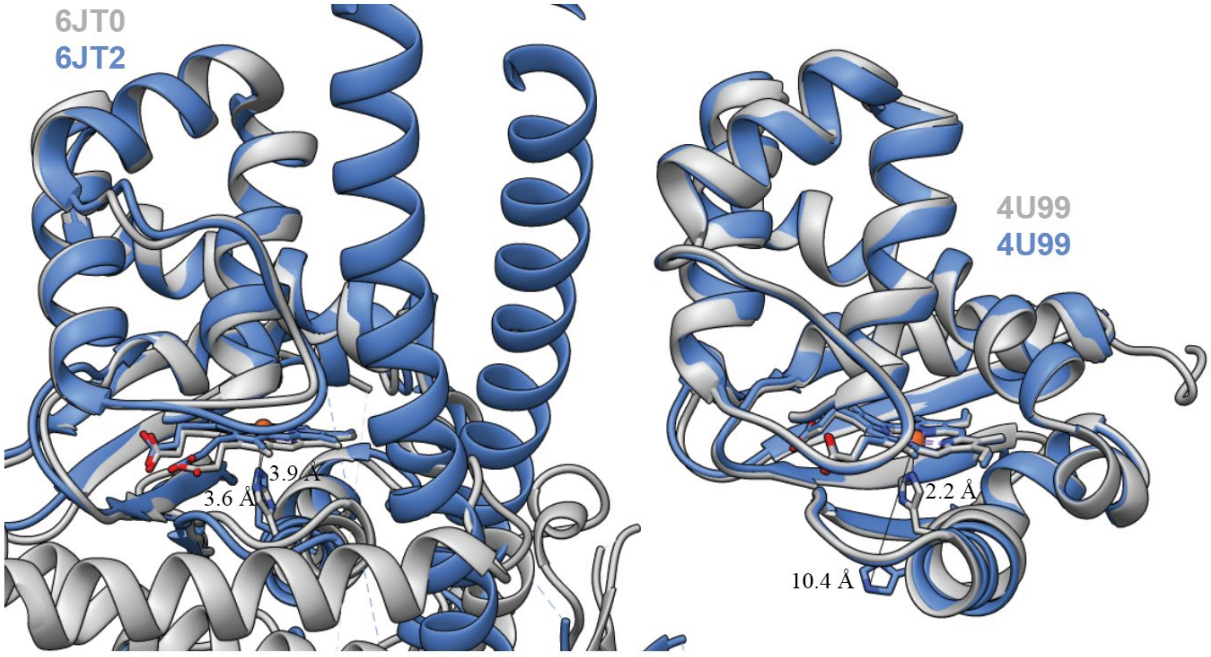


Figure 5.6 – Canonical motions of the H-NOX domain. Comparison of the rotations of the proximal histidine in *Hs* sGC and *So* H-NOX structures in the unliganded (gray) and NO-bound (blue) state. Electron density shows that the proximal histidine is cleaved from the iron in both cases, however in the bacterial H-NOX the αF helix (on which the proximal histidine is located) rotates ~90° where a similar rotation is not observed in the *Hs* sGC structures.

There has been some investigation into other residues on the αF helix of the β H-NOX domain that are involved in the initiation of signal transduction from the βH-NOX to the catalytic domains, namely βD102 and βT110.[13,14] The distance of βD102 ε nitrogen to βR258 decreases from 10.9 Å in the inactive state to 3.0 Å in the active state.[9] These residues, along with βR40, βY83, βH105 (proximal histidine), βD106 (D/E strictly conserved) (highlighted in Kang et. al)[9], βH107, βT110, βY112, βS122, βH260, βR305 (R/K), βK307, βY363, βT366 and βE370 on the β subunit along with αD411, αQ418, and αQ422 of the α subunit are implicated in the β H-NOX interface with the other domains of sGC in the active state (Figure 5.8). Both conservative mutations and radical mutations of these residues, coupled to rescue double mutants along this interface are warranted to dissect the activation pathway of sGC.

The coiled coil domains of sGC are now understood to act as a hinge that transduces the signal between the gas-sensing and catalytic domains of sGC. In the inactive state, each coiled coil domain is broken into two coiled coil regions connected by linkers αA419-αK426 and βL354-

βF359. Kang et al. made the radical mutations αD423P and βG356P and showed that introducing kinks in the coiled coil domains prevent activation.[9] Different types of mutations in this area can expand our understanding of the activation pathway in this region. Specifically, the introduction of (GSG)n linkers where n=1 or 2 at these sites could uncouple conformational changes in the β H-NOX from the catalytic domains. Mutations that more effectively zip up the coiled-coil by introducing a leucine zipper into the inactive conformation of sGC might produce sGC that is constitutively activated or maximally active with only one equivalent of NO. Examples of such mutations could be combinations of α(A420L/Q423L), α(A420L/D434L), β(Q349L/D353L), β(Q349L/D353L), and triple mutants that also include αQ423L and βF350L.

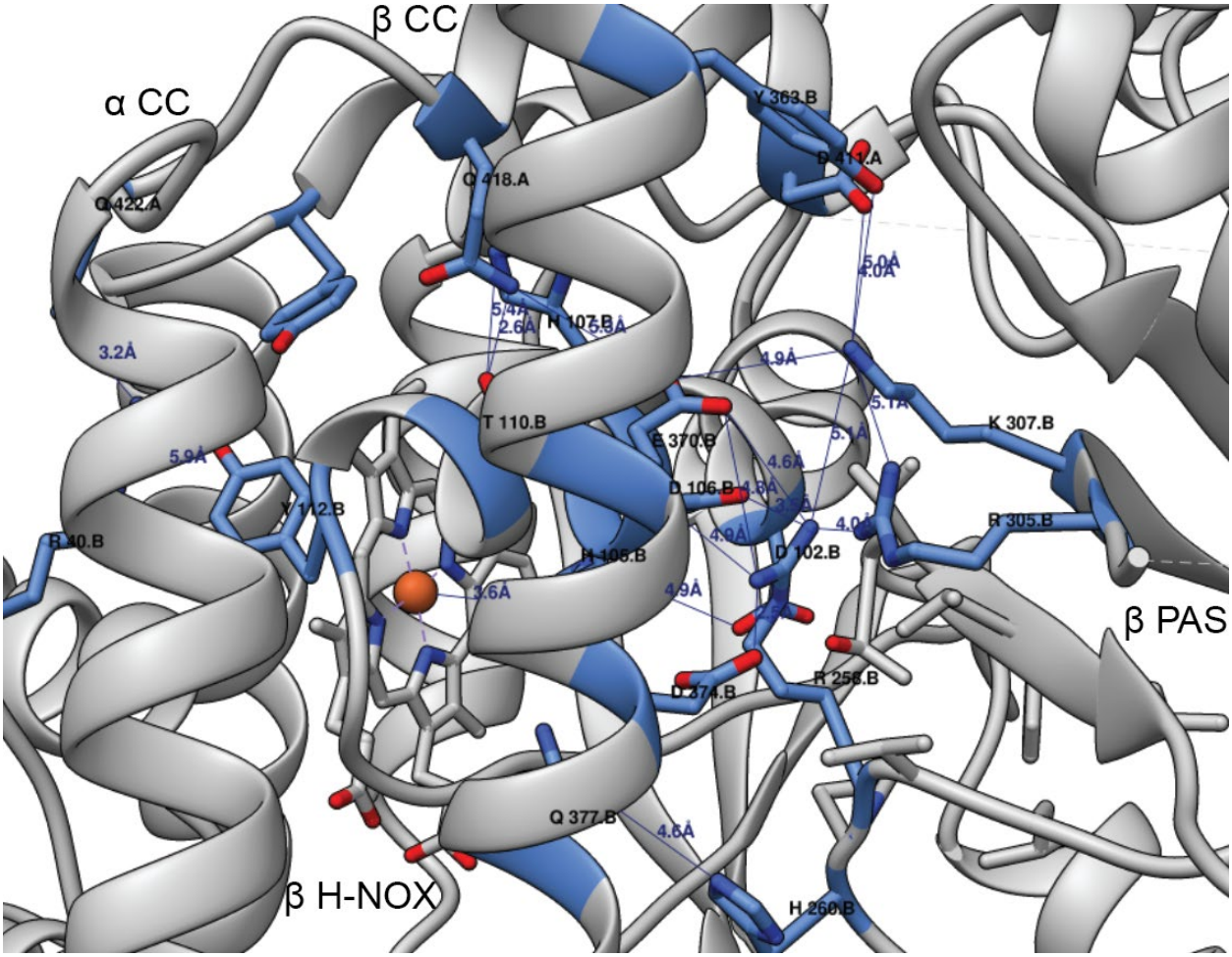


Figure 5.7 - The inactive state polar contacts between the β H-NOX domain and the β PAS and α CC domain.

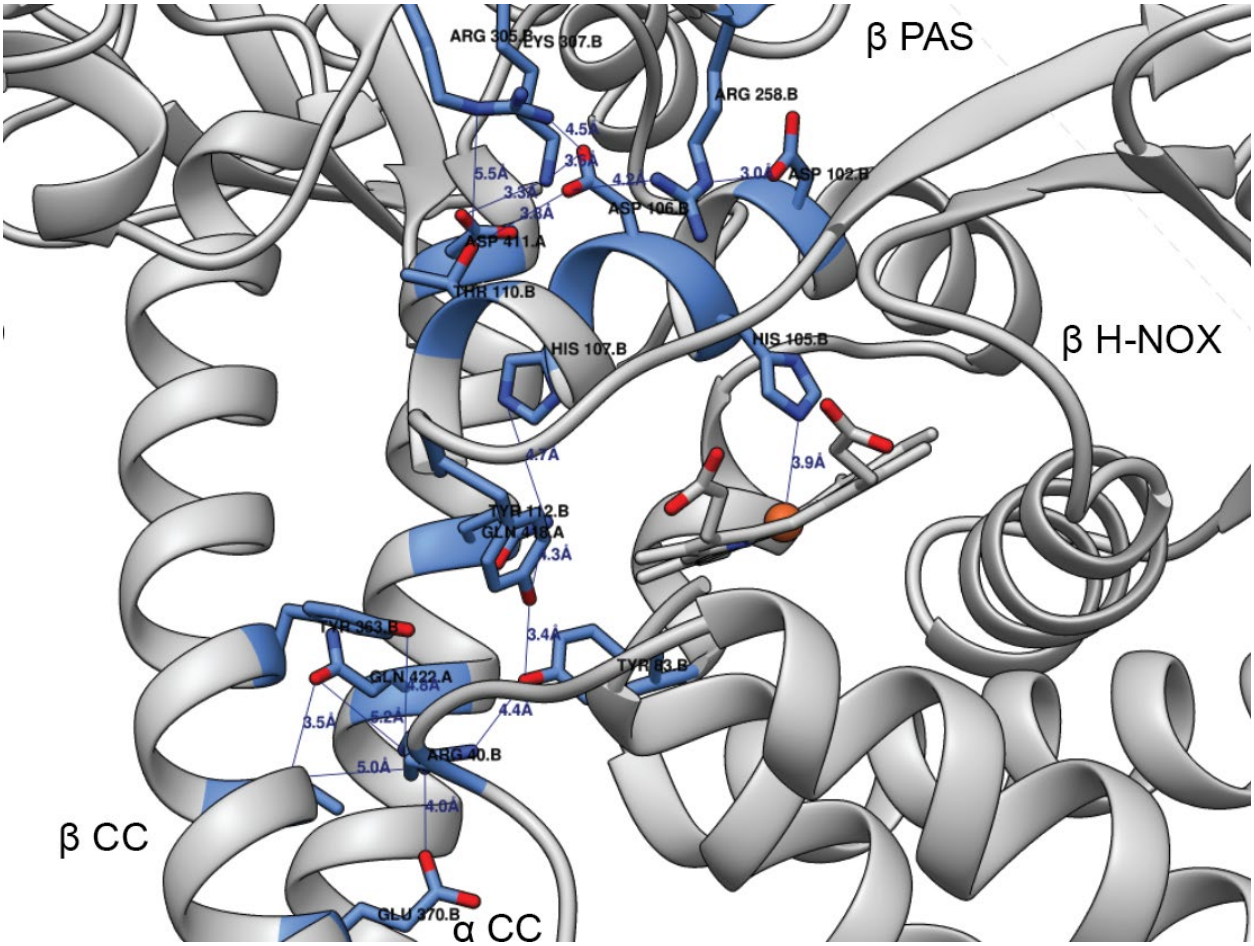


Figure 5.8 - The active state polar contacts between the β H-NOX domain and the β PAS and α CC domain.

The identification of the putative binding pocket on sGC of YC-1 discussed in Chapter 2 is also worthy of mutational analysis. Combining crystal structure 6JT2 and the EM map of *Manduca sexta* with NO and YC-1 allows for identification of residues that are within 5 Å of the proposed binding site (Figure 5.9). Residues close to this volume include βF4, βV39, βR40, βF77, βS81, βY83, βT112, βT366, and βE370 on the β subunit and αQ422, αL425, αR428, and αL429 on the α subunit. Conservative or radical mutations of these residues on sGC truncations could be used to look at the effect of YC-1 on CO binding the YC-1 dissociation constant[15,16], or the EC_{50} of YC-1 for activation of full-length sGC in the 1-NO state. Radical mutations that do not interrupt full activation by excess NO but do interrupt full activation of the 1-NO state by YC-1 would be indicative of a disrupted binding pocket. To see if this site is indeed the only site for small molecule drug binding to sGC, cryo-EM of sGC in the presence of NO and stimulators from Bayer and Cyclerion Therapeutics (formally Ironwood Inc.). If these compounds indeed bind in the same pocket, this would lend more confidence to the hypothesis that stabilization of the β H-NOX domain against the CC domains is a similar mechanism of activation for all classes of sGC stimulators.

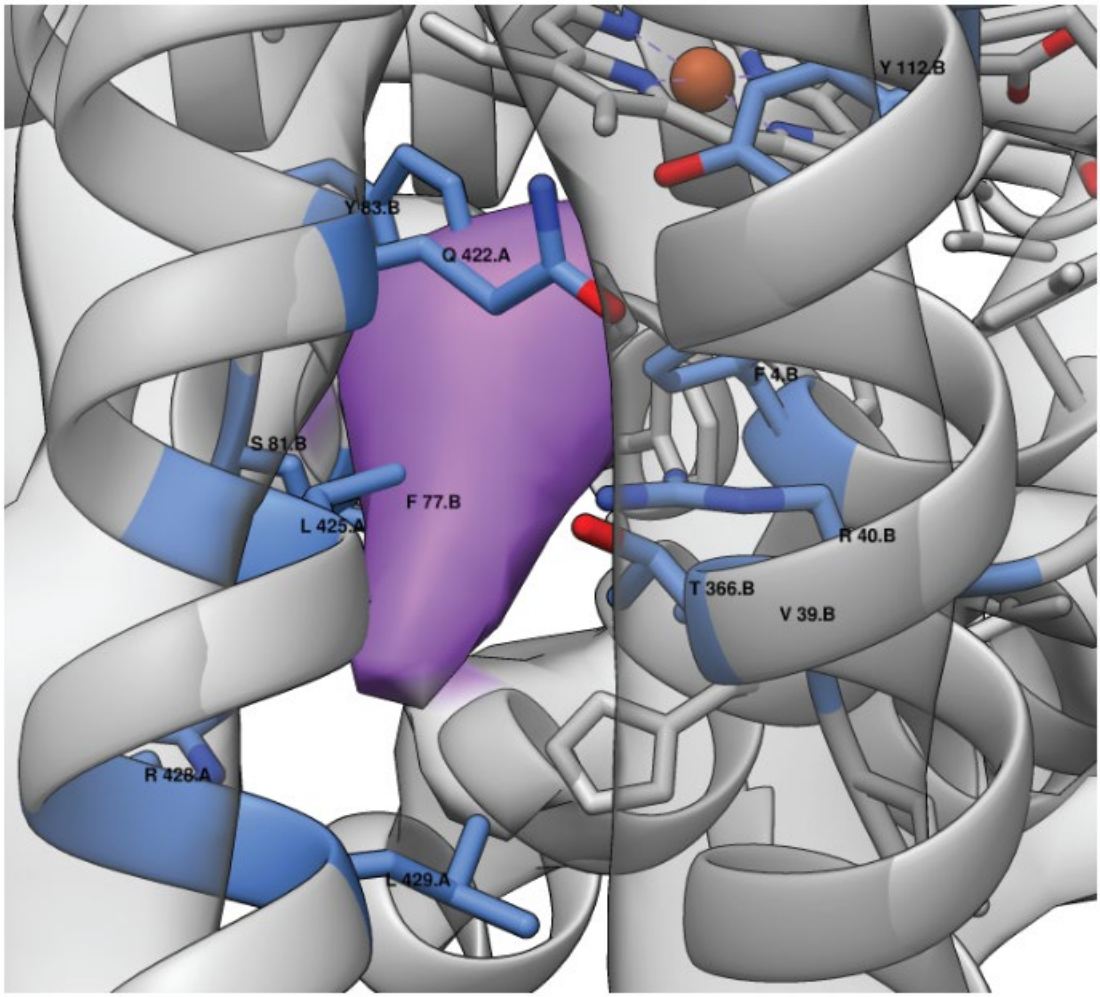


Figure 5.9 – Proposed YC-1 binding pocket and interacting residues. The proposed YC-1 binding pocket is shown in purple. Residues that surround this pocket are colored in blue.

There is overlap of some residues that seem to be involved in the interfaces between the β H-NOX domain and other domains of sGC and are also close in space to the putative YC-1 binding site. Specifically, βR40, βY83, βT112, βT366, and βE370 on the β subunit and αQ422 on the α subunit seem to be critical for both the contacts the β H-NOX subunit makes with other domains and are close to the YC-1 binding pocket. Great care should be exercised when analyzing activity data from variants with these specific residues mutate so as to not conflate β H-NOX interdomain interactions for YC-1 activation. One alternate strategy that uses orthogonal methods to investigate the YC-1 binding site would be to use Förster Resonance Energy Transfer (FRET) between a site on the protein and a modified stimulator. By selecting emission and excitation wavelengths accordingly, it should be possible to attach one fluorophore to sGC at known location (e.g. substituting the heme with a fluorescent analog, using a fluorescent nucleotide analog, attaching a fluorophore to the C-terminus of one of the catalytic domains) and using a sGC stimulator modified with an appropriate donor or acceptor, the distance could be calculated. If two of these approaches

(or alternate approaches) succeed, the intersection of the diameters of the spheres away from the known site should agree with the location of YC-1 in the structure.

Experiments described in Chapter 2 used SEC-SAXS to observe a partial elongation of sGC in the 1-NO state. As new improvements are constantly being made to the system up at LBNL, future scattering experiments are warranted. Firstly, changes in the experimental protocol to ensure the concentration of NO is maximal in the sample flow cell should be taken, either by increasing the amount of NO donors used or through the use of a new flow cell that can add fast-releasing donors directly as the protein is passing through the flow cell. The initial impetus for these experiments was the promise of time resolved SAXS, which would require the use of photocaged NO donors, or perhaps a timed release of NO donor with a new design of flow cell. The ability to measure the kinetics of the conformational change upon activation of NO would provide conformation of more indirect measurements of this structural rearrangement.

NO causes striking conformational changes in sGC. Now that we have the endpoints of this activation, it is possible to ask questions about the energetics of this motion. There are three likely scenarios as to what the free energy landscape looks like in terms of sGC conformational state. The first is of a low free energy inactive state and a high free energy active state, similar to that of a spring stretched past its point of equilibrium (sGC is being "stretched" by NO). The second is of a high free energy inactive state and a low free energy active state, reminiscent of a compressed spring that is constrained in that position (with excess NO "unlocking" the restraint). The final scenario involves the lowest free energy conformation changing based on NO concentration, similar to a spring made of copper wire that can be deformed. At this point, scenarios one and three seem the most plausible because it seems necessary for a sufficient driving force to be present for sGC to return to the 1-NO state after removal of excess NO. However, it would be interesting to examine these three possibilities using single molecule experiments and capture the forces of the motion of sGC in the presence of different stimulators.[17] To accomplish this, sGC would need to be suspended between two objects, either two beads with optical tweezers or a cantilever and a piezoelectric scanner for atomic force microscopy. It has been relatively easy to add tags onto the C-termini of sGC; sGC is currently purified with a His_6 tag added to the end of the α catalytic domain. However, attaching molecules to the N-terminus of sGC would be more challenging because the N-terminus of the β H-NOX is buried. One possibility could be to use N-terminal chemistry (e.g. 2-PCA) and screen for conditions where the α H-NOX is modified but the β H-NOX is left unmodified.[18] Site-specific unnatural amino acid incorporation could also be examined, however with the challenges already associated with sGC expression this might not be feasible. These experiments would allow for the exploration of both the thermodynamics and kinetics of sGC conformational change and would provide additional insight in to the activation mechanism of the canonical gas-receptor protein.

As discussed in Chapter 2, work in this thesis and the structure by Kang et al are the first two examples of full-length structures in both inactive and active states that contain the regulatory domain known as the S-helix.[19] (Note: the proteins with known full-length structures described by Montfort et al to contain S-helix domains do not possess the typical sequence motif as explained by Aravind et al).[19,20] Proteins that contain the S-helix motif possess both sensing domains and output domains that are connected by the S-helix, which is thought to communicate the presence of signaling molecule in one domain to catalytic output in the other. In sGC, the S-helix motif

connects the gas-sensing N-terminal regulatory lobe from the C-terminal catalytic domains and is contained within the C-terminal portion of the coiled-coil domains (residues αK427-I457 and βR360-L394). However, it is not clear whether the bend in the helices that was observed in sGC is a feature of all proteins that contain an S-helix or if the general function of the S-helix results in the twisting of the coiled-coils. Structural studies of other proteins that contain S-helix domains could provide insight into this conserved signaling strategy employed by receptor proteins. For example, the cyclic di-GMP phosphodiesterase PA4781 (Uniprot: Q9HV27) from *Pseudomonas aeruginosa* contains an S-helix that connects its REC domain from the HD-GYP domain, but only truncations of protein have been solved.[21] Another example is the transcriptional regulator NifA (Uniprot: O66591) from *Aquifex aeolicus (strain VF5)* represents a different class of regulator in which only truncations of the protein have been solved.[22] Similar to sGC, cryo-EM could be used in this case to obtain the full-length structure of these receptors in the on and off state. There are possibly many more examples of soluble proteins with S-helix domains, however currently there are no structures of any individual domains; a large portion of these proteins do have transmembrane domains which, while presenting additional challenges, can still be interrogated structurally.[23] These structures could then be used to compare S-helix domains across two different types of signaling proteins to understand the commonalities furnished by the S-helix and which aspects are unique to each receptor.

Pertaining to Chapter 3 and the search for the second NO binding site on sGC, labeling sGC with titrations of bioconjugational agents and testing for both average mass shift as well as activity could be particularly insightful to see if at a certain average number of amino acids labeled, sGC always loses activity. Perhaps this is specific to cysteines, or perhaps with more lysines or carboxylic acid residues labeled sGC might always lose the activity with excess NO. If it turns out that the same average labeling of sGC with different bioconjugational agents yields different activity profiles, using mass spectrometry to sequence peptides in each condition could potentially yield a pattern for amino acids modified in conditions that always loose activity or always maintain activity. Additionally, other disulfide exchange reagents with varying steric bulk could be tested to see if the size of MMTS is critical in its ability to prevent activation by excess NO. For example, 2-hydroxyethyl disulfide , 3-carboxylethyl disulfide, and 5,5'-dithiobis-(2-nitrobenzoic acid) (which would also provide a colorimetric readout for how many cysteines modified) would constitute a steric size series to test if smaller disulfide exchange reagents inhibit sGC to a larger extent. In a slightly different vein, the double addition type experiments where first IAM is added followed by addition of MMTS or another addition of IAM should be followed with high resolution mass spectrometry. Any cysteine labeled with IAM would be considered to be not critical, but any cysteine seen with a MMTS labeled would be of interest. For these types of mass spectrometry experiments, the observation of a non-labeled cysteine does not mean that it was never labeled. It is only the observation of a labeled peptide that can be commented on. With the full-length structures now published, any proteomic data can then be overlaid on the structures to attempt to rationalize these results.

A different line of research that should be undertaken is the potential modification of βC541, the residue in the active site that discriminates between guanylate cyclases and adenylate cyclases (Asp in adenylate cyclases). If this residue is modified by MMTS or NEM, this could explain the lack of activity of sGC labeled with these molecules. Either a full kinetic analysis should be conducted with sGC in the excess NO state with and without MMTS to compare K_M

values, or some type of biophysical measurement of the K_D (equilibrium dialysis, ITC, or fluorescence with a substrate analogue) will be an important control to ensure that the binding of the nucleotide is not impaired in the labeling of sGC. If βC541 is being modified by cysteine alkylating agents, that would greatly complicate the interpretation of the biochemical activity assays.

Work presented in Chapter 2 has shown that conformational changes of sGC can be visualized with biophysical methods such as cryo-electron microscopy and small angle X-ray scattering. Comparison of electron densities of sGC with excess NO to sGC labeled with MMTS with excess NO could help decipher whether or not MMTS is preventing a conformational change. Additionally, comparison of SAXS scattering profiles of sGC with excess NO to sGC labeled with MMTS with excess NO could yield subtle differences that could indicate whether or not MMTS is indeed preventing a conformational change.

The slight conformational changes in the β H-NOX domain could potentially drive the large-scale conformational changes seen in active sGC. Kang et al describe the active β H-NOX domain conformation as being incompatible with the rest of the inactive structure.[9] One hypothesis that is worth exploring is the ability of MMTS to inhibit this small conformational change in the in the β H-NOX. If MMTS is labeling cysteines in the β H-NOX domain that prevent the domain from adopting its active conformation, perhaps MMTS is not blocking a second site cysteine at all. There are two conserved cysteines in the β H-NOX domain, βC78 and βC122. It is possible that variants of these cysteines could render sGC tolerant to MMTS labeling. Past work from other labs and work in our lab has shown that the sGC variant βC78S does not bind heme, however serine is more polar than cysteine.[24] Proposed variants of sGC to explore this hypothesis would be the double mutant βC78A/V in combination with βC122A/V. If this variant has an activation profile similar to WT labeled with MMTS, then different cysteines are involved. However, if this variant has an activation profile similar to WT not labeled with MMTS, this could be the mechanism behind MMTS inhibition of sGC.

Research presented in Chapter 4 and this chapter investigated one of the six sGC homologues found in the eukaryotic alga *Chlamydomonas reinhardtii*, however interesting questions still remain in this area. To begin, it is not clear why a Cyg11 homodimer with two heme binding sites only binds one heme cofactor. The current hypothesis is that Cyg11 adopts some sort of asymmetric conformation so as to allow the binding of one heme but not the other. Biophysical experiments such as SAXS and cryo-EM could be used to test this hypothesis. Additionally, the final two (*cyg56* and *cyg57*) sGCs were challenging to clone, but perhaps the purchase of codon optimized vectors would allow for the biochemical investigation of these final homologues. This chapter began to investigate targeted transcriptomics of all of *C. reinhardtii*'s sGCs, however these efforts should be repeated. Finally, although *C. reinhardtii* is not the most genetically tractable model organism, there exists a transposon insertional mutant library that potentially have these sGCs knocked out.[25] In collaboration with the Merchant lab, mutants with insertions in the CDS of *Cr* sGCs should be examined phenotypically and transcriptomically in regular and stress growth conditions (including lack of metals, lack of nitrogen, lack of light, etc.).

Finally, new work has been recently published that describes a new species of choanoflagellate that forms disc-shaped colonies which change their convexity based on light or

mechanical stimuli.[26] This process is proposed to be regulated by a rhodopsin fused to a cGMP-specific phosphodiesterase, but the upstream signaling of this process is not yet understood. Intriguingly, this organism possesses genes homologous to mammalian nitric oxide synthase (NOS) proteins and sGCs. Preliminary data indicate that NO donors induce convexity changes, strengthening the argument that NOS and sGC play a role in the communication of this choanoflagellate colony. NOSs have been successfully expressed in *E. coli* previously, but if the choanoflagellate sGC is responsive to NO and can be expressed in *E. coli* that would be very intriguing. Additionally, it would be interesting to see if this sGC is responsive to NO and has the three-state activation model similar to mammalian sGCs.

References

(1) de Montaigu, A.; Sanz-Luque, E.; Galván, A.; Fernández, E. A Soluble Guanylate Cyclase Mediates Negative Signaling by Ammonium on Expression of Nitrate Reductase in *Chlamydomonas*. *Plant Cell* **2010**, *22* (5), 1532–1548.

(2) Horst, B. G.; Stewart, E. M.; Nazarian, A. A.; Marletta, M. A. Characterization of a Carbon Monoxide-Activated Soluble Guanylate Cyclase from Chlamydomonas Reinhardtii. *Biochemistry* **2019**, *58* (17), 2250–2259.

(3) Zhao, Y.; Brandish, P.; Ballou, D.; Marletta, M. A. A Molecular Basis for Nitric Oxide Sensing by Soluble Guanylate Cyclase. *Proc. Natl. Acad. Sci. U. S. A.* **1999**, *96* (26), 14753–14758.

(4) Morton, D. B.; Langlais, K. K.; Stewart, J. a; Vermehren, A. Comparison of the Properties of the Five Soluble Guanylyl Cyclase Subunits in Drosophila Melanogaster. *J. insect Sci.* **2005**, *5* (12), 1–10.

(5) Allorent, G.; Lefebvre-legendre, L.; Chappuis, R.; Kuntz, M.; Truong, T. B.; Niyogi, K. K.; Ulm, R.; Goldschmidt-Clermont, M. UV-B Photoreceptor-Mediated Protection of the Photosynthetic Machinery in *Chlamydomonas Reinhardtii*. *PNAS* **2016**, *113* (51), 14864–14869.

(6) Yoo, B.-K.; Lamarre, I.; Martin, J.-L.; Rappaport, F.; Negrerie, M. Motion of Proximal Histidine and Structural Allosteric Transition in Soluble Guanylate Cyclase. *Proc. Natl. Acad. Sci.* **2015**, *112* (14), E1697–E1704.

(7) Urzica, E. I.; Casero, D.; Yamasaki, H.; Hsieh, S. I.; Adler, L. N.; Karpowicz, S. J.; Blaby-Haas, C. E.; Clarke, S. G.; Loo, J. A.; Pellegrini, M.; et al. Systems and *Trans*-System Level Analysis Identifies Conserved Iron Deficiency Responses in the Plant Lineage. *Plant Cell* **2012**, *24* (10), 3921–3948.

(8) Castruita, M.; Casero, D.; Karpowicz, S. J.; Kropat, J.; Vieler, A.; Hsieh, S. I.; Yan, W.; Cokus, S.; Loo, J. a; Benning, C.; et al. Systems Biology Approach in Chlamydomonas Reveals Connections between Copper Nutrition and Multiple Metabolic Steps. *Plant Cell* **2011**, *23* (4), 1273–1292.

(9) Kang, Y.; Liu, R.; Wu, J.-X.; Chen, L. Structural Insights into the Mechanism of Human Soluble Guanylate Cyclase. *Nature* **2019**.

(10) Herzik, M. A.; Jonnalagadda, R.; Kuriyan, J.; Marletta, M. A. Structural Insights into the Role of Iron–Histidine Bond Cleavage in Nitric Oxide-Induced Activation of H-NOX Gas Sensor Proteins. *Proc. Natl. Acad. Sci.* **2014**, *111* (40), E4156–E4164.

(11) Hespen, C. W.; Bruegger, J. J.; Phillips-Piro, C. M.; Marletta, M. A. Structural and Functional Evidence Indicates Selective Oxygen Signaling in Caldanaerobacter Subterraneus H-NOX. *ACS Chem. Biol.* **2016**, *11* (8), 2337–2346.

(12) Smith, B. C.; Underbakke, E. S.; Kulp, D. W.; Schief, W. R.; Marletta, M. a. Nitric Oxide Synthase Domain Interfaces Regulate Electron Transfer and Calmodulin Activation. *Proc. Natl. Acad. Sci. U. S. A.* **2013**, *110* (38), E3577-86.

(13) Baskaran, P.; Heckler, E. J.; Van Den Akker, F.; Beuve, A. Aspartate 102 in the Heme Domain of Soluble Guanylyl Cyclase Has a Key Role in No Activation. *Biochemistry* **2011**, *50* (20), 4291–4297.

(14) Baskaran, P.; Heckler, E. J.; van den Akker, F.; Beuve, A. Identification of Residues in the Heme Domain of Soluble Guanylyl Cyclase That Are Important for Basal and Stimulated Catalytic Activity. *PLoS One* **2011**, *6* (11), e26976.

(15) Fritz, B.; Roberts, S.; Ahmed, A.; Breci, L.; Li, W.; Weichsel, A.; Brailey, J.; Wysocki, V.; Tama, F.; Montfort, W. Molecular Model of a Soluble Guanylyl Cyclase Fragment Determined by Small-Angle X-Ray Scattering and Chemical Cross-Linking. *Biochemistry* **2013**, *52* (9), 1568–1582.
(16) Wales, J. A.; Chen, C.; Breci, L.; Weichsel, A.; Bernier, S. G.; Sheppeck II, J. E.; Solinga, R.; Nakai, T.; Renhowe, P. A.; Jung, J.; et al. Discovery of Stimulator Binding to a Conserved Pocket in the Heme Domain of Soluble Guanylyl Cyclase. *J. Biol. Chem.* **2018**, *293* (5), 1850–1864.
(17) Neuman, K. C.; Nagy, A. Single-Molecule Force Spectroscopy: Optical Tweezers, Magnetic Tweezers and Atomic Force Microscopy. *Nat. Methods* **2008**, *5* (6), 491–505.
(18) Macdonald, J. I.; Munch, H. K.; Moore, T.; Francis, M. B. One-Step Site-Specific Modification of Native Proteins with 2-Pyridinecarboxyaldehydes. *Nat. Chem. Biol.* **2015**, *11* (5), 326–331.
(19) Anantharaman, V.; Balaji, S.; Aravind, L. The Signaling Helix: A Common Functional Theme in Diverse Signaling Proteins. *Biol. Direct* **2006**, *1*, 25.
(20) Montfort, W. R.; Wales, J.; Weichsel, A. Structure and Activation of Soluble Guanylyl Cyclase, the Nitric Oxide Sensor. *Antioxidants redox Signal.* **2016**, *00* (00), 1–16.
(21) Rinaldo, S.; Paiardini, A.; Stelitano, V.; Brunotti, P.; Cervoni, L.; Fernicola, S.; Protano, C.; Vitali, M.; Cutruzzolà, F.; Giardina, G. Structural Basis of Functional Diversification of the HD-GYP Domain Revealed by the Pseudomonas Aeruginosa PA4781 Protein, Which Displays an Unselective Bimetallic Binding Site. *J. Bacteriol.* **2015**, *197* (8), 1525–1535.
(22) Batchelor, J. D.; Lee, P. S.; Wang, A. C.; Doucleff, M.; Wemmer, D. E. Structural Mechanism of GAF-Regulated Σ54 Activators from Aquifex Aeolicus. *J. Mol. Biol.* **2013**, *425* (1), 156–170.
(23) Qi, C.; Sorrentino, S.; Medalia, O.; Korkhov, V. M. The Structure of a Membrane Adenylyl Cyclase Bound to an Activated Stimulatory G Protein. *Science (80-.).* **2019**, *364* (6438), 389–394.
(24) Friebe, A.; Wedel, B.; Harteneck, C.; Foerster, J.; Schultz, G.; Koesling, D. Functions of Conserved Cysteines of Soluble Guanylyl Cyclase. *Biochemistry* **1997**, *36* (6), 1194–1198.
(25) Li, X.; Patena, W.; Fauser, F.; Jinkerson, R. E.; Saroussi, S.; Meyer, M. T.; Ivanova, N.; Robertson, J. M.; Yue, R.; Zhang, R.; et al. A Genome-Wide Algal Mutant Library and Functional Screen Identifies Genes Required for Eukaryotic Photosynthesis. *Nat. Genet.* **2019**, *51* (4), 627–635.
(26) Brunet, T.; Larson, B. T.; Linden, T. A.; Vermeij, M. J. A.; McDonald, K.; King, N. Light-Regulated Collective Contractility in a Multicellular Choanoflagellate. *Science (80-.).* **2019**, *366*, 326–334.

APPENDICES

Appendix A

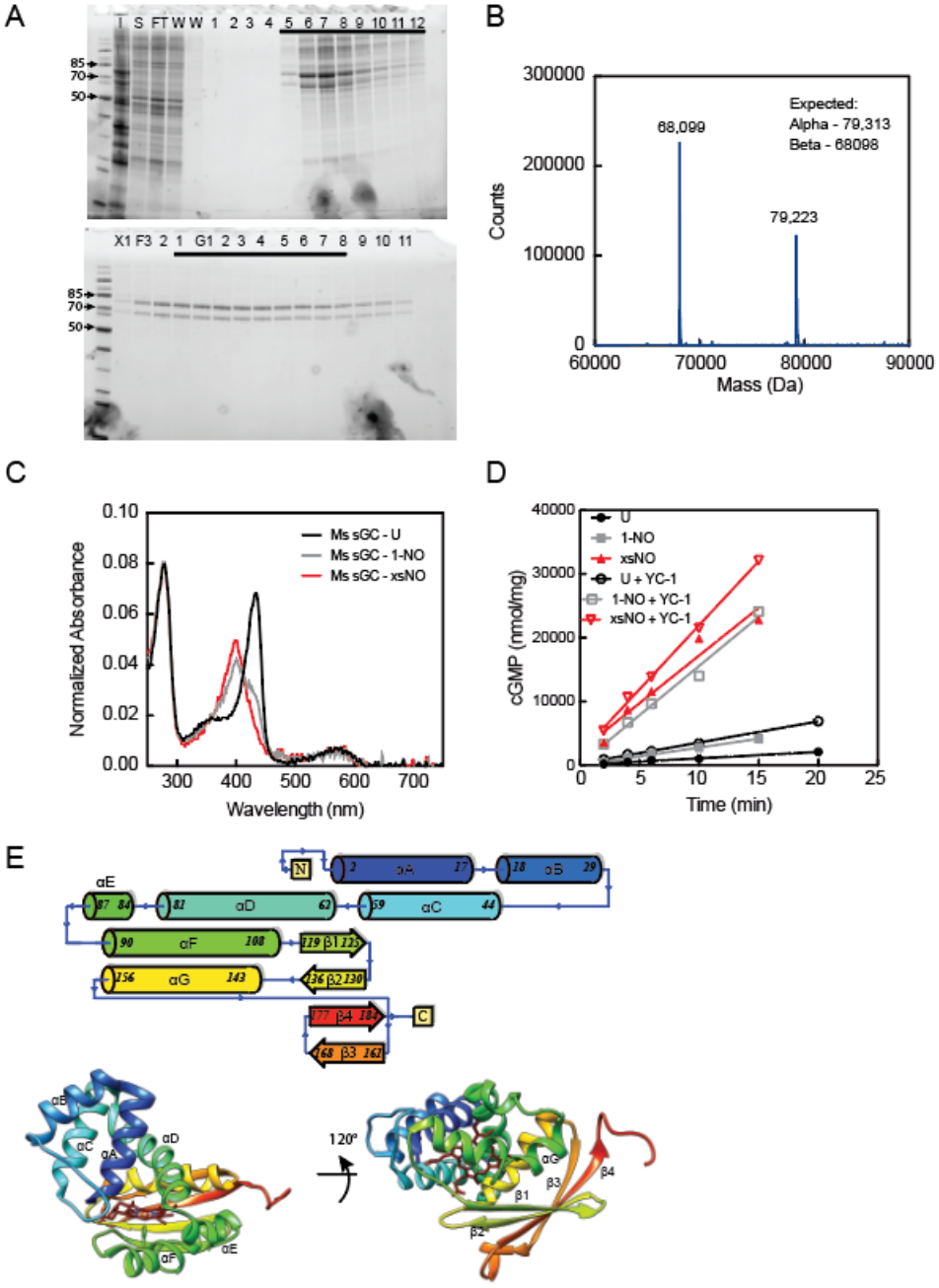


Figure A.1 – Biochemical characterization of full-length *Ms* sGC. A) Representative SDS-PAGE gels for purification of *Ms* sGC. B) Deconvoluted intact protein mass spectrum of purified *Ms* sGC. C) UV–visible absorption spectra of *Ms* sGC used in activity assays. The 1-NO sample was generated by first adding excess NO followed by buffer exchange. D) Representative time courses of the cGMP assay for *Ms* sGC in different ligation states. The cGMP product was quantified using an enzyme linked immunosorbent assay. E) Topology diagram of β H-NOX secondary structure. Bottom Left: Front view of the β H-NOX displaying helices αA–αF. Bottom Right: Bottom view of the β H-NOX displaying helix αG and sheets β1–β4.

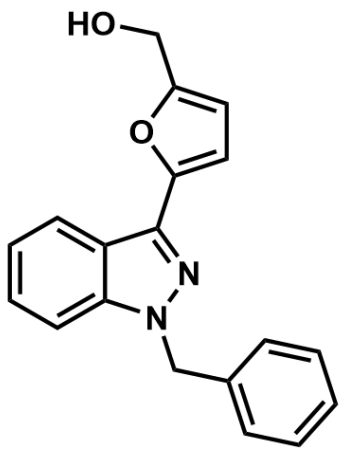


Figure A.2 – The structure of the benzyl imidazole compound YC-1.

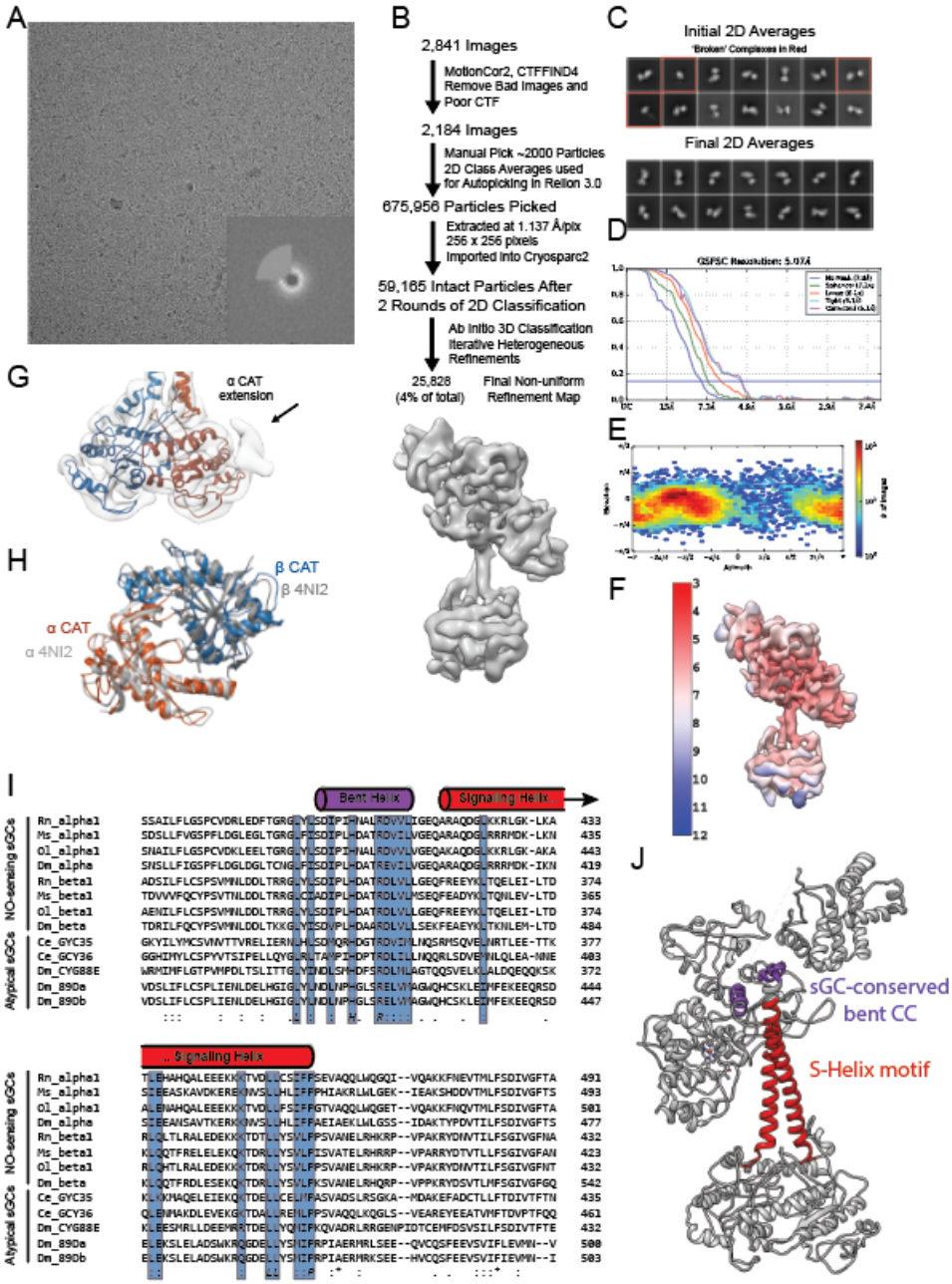


Figure A.3 – Cryo-EM of the *Ms* sGC inactive state and S-helix alignments A) Representative micrograph with inlet of corresponding FFT. B) Electron microscopy flowchart for processing of inactive sGC data with three views of the final reconstruction shown colored in gray. C) Selected 2D class averages. D) FSC curve from cryosparc2 with gold standard 0.143 FSC shown in blue. E) Plot of single particle population based on orientation parameters. F) Inactive reconstruction colored by local resolution calculated using cryosaprc2 implementation of blocres. G) Close up view of the modeled inactive CAT dimer highlighting unmodeled density near the α CAT C-terminus. H). Overlay of the PDB: 4NI2 (grey) and the inactive CAT dimer (colored by α and β). I) The coiled-coil domain of eight sequences from four heterodimeric biochemically-verified NO-responsive sGCs and five sequences from atypical sGCs that sense O_2 are aligned, with conserved residues are highlighted in blue. The approximate domain architecture from the inactive state CC domains are shown above the sequences. J) Inactive model of *Ms* sGC is shown with the bent helix in purple and the S-helix in red.

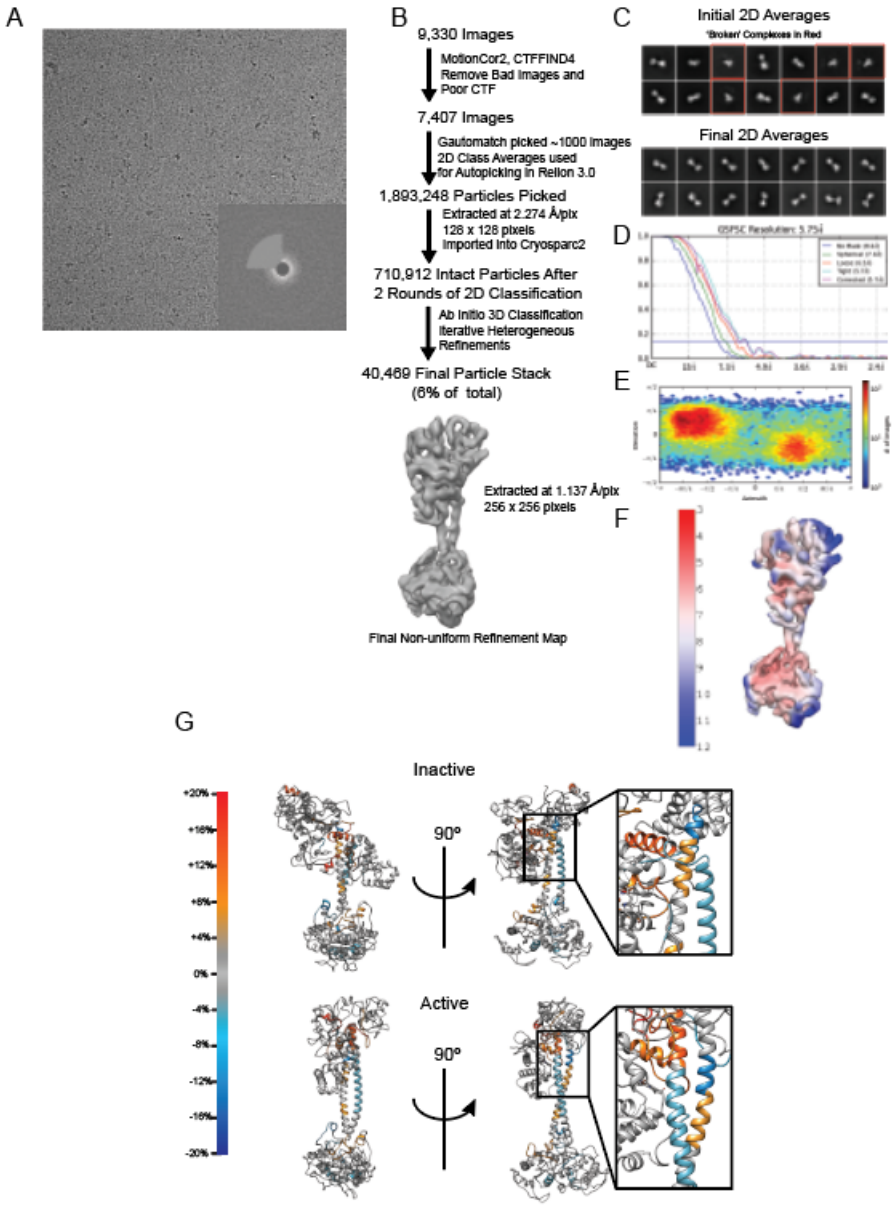


Figure A.4 – Cryo-EM of the *Ms* sGC active state and HDX-MS overlay A) Representative micrograph with inlet of corresponding FFT.B) Electron microscopy flowchart for processing of active sGC data.C) Selected 2D class averages.D) FSC curve from cryosparc2 with gold standard 0.143 FSC shown in blue.E) Plot of single particle population based on orientation parameters.F) Active reconstruction colored by local resolution calculated using cryosaprc2 implementation of blocres. G) Differential H/D Exchange values representing change in percent D by subtracting the active state from the inactive state. Regions with large changes in percent D overlay with secondary structural elements observed in the structures to undergo large conformational changes, including the αF helix of the β H-NOX, the bent CC region, and S-helix motif, and certain helices of the CAT domain. The insets show a close up view of the H/D exchange patters of the CC domains.

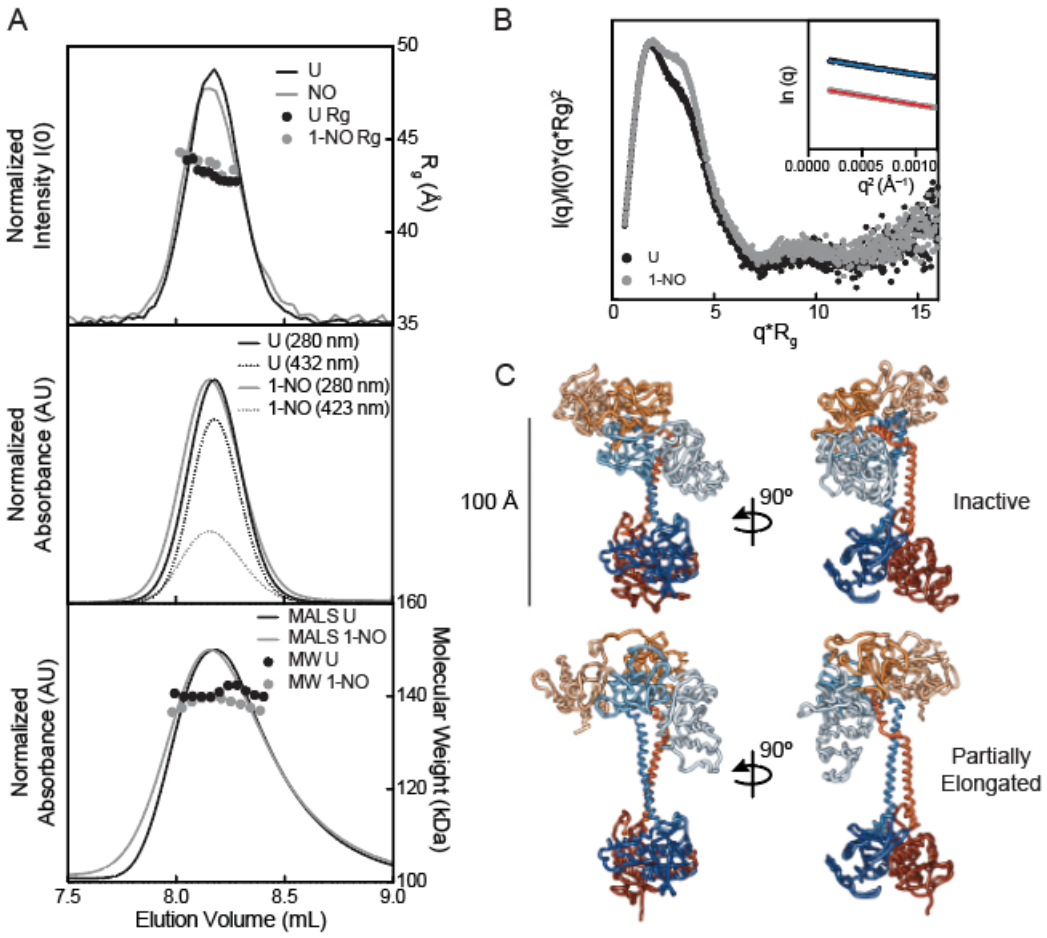


Figure A.5 – SAXS of *Ms* sGC. A) Top: Small angle X-Ray scattering chromatograms with intensity (lines) and R_g (symbols) values for each frame merged across the SEC peak. Middle: UV-visible absorbance chromatograms showing the normalized absorbance at 280 (line) and 432 nm (dotted-line). Bottom: Multi angle light scattering chromatograms with light scattering signal (line) and molecular weight (dots). All chromatograms aligned to the MALS peaks which are furthest downstream to compensate for variations in timing and band broadening. B) Normalized Kratky plot for the inactive (black) and active state (gray) of sGC. Inset: Guinier plots within the q*Rg < 1.3. C) Top: Best fit conformations for the inactive state of sGC. Bottom: Extended structure from the best ensemble for the activated sGC conformation

Table A.1
Cryo-EM data acquisition, image processing and model refinement

Imaging/Processing	Inactive	Active
Microscope	Talos Arctica	Talos Arctica
Voltage (kV)	200	200
Camera	K3	K3
Defocus range (μm)	-1.5 to -3.5	-1.5 to -3.5
Pixel size (Å)	1.137	1.137
Total electron dose (e-/Å^2)	60	60
Exposure time (s)	6	6
Number of movies	2,841	9,330
Number of frames/movie	60	60
Initial particle number	675,956	710,912
Final particle number	25,828	40,469
FSC 0.143 (unmasked/masked, Å)	8.3 / 5.1	7.6 / 5.8
Refinement		
Homology Models (PDB)	α/β H-NOX (2O0C), α/β PAS (4GJ4), α/β CC (3HLS), α CAT (3UVJ), β CAT (2WZ1)	
Cα Residues	1129	953
Ligands	Heme	Heme
RMSDs		
Bond lengths (Å)	0.006	0.009
Bond angles (°)	1.096	1.229
Ramachandran		
Favored (%)	82.57	76.35
Allowed (%)	17.25	23.23
Outlier (%)	0.09	0.42
All Atom Clash score	1.38	6.92
Molprobity score	1.57	2.16
Accession codes		
EMDB	EMD-20282	EMD-20283
PDB	PDB - 6PAS	PDB – 6PAT

Table A.2
SEC-SAXS-MALS-UV-visible absorption results for the activation of sGC

Measurement	Inactive	1-NO
SAXS Molecular Weight (kDa)	~140	~131
MALS Molecular Weight (kDa)	140.7 (±0.008%)	138.4 (±0.014%)
Ratio 432/280 nm Absorbance	0.765	0.319
Radius of Gyration, R_g (Å)	43.1 (±0.4)	43.8 (±0.2)
Maximum Dimension, D_{max} (Å)	133	142
Porod Exponent, P_x	4	4
Inactive Conformation (%)	100	72
Partially Extended Conformation (%)	0	28
χ^2-value of Fit	1.44	1.51

Appendix B

S1A

```
Cr_Cyg11      MIGWINISVESFVREKFGDAVWEQTLSTSGVQAGWVSSCPYPDAATYGLVISASNILGITSAQVL
              EAFGVYFVEYTARLGYEKLLKSLGSNMAEFLKNLNDLHLHLSMSFPSMAAPAFKCTDVGPTCLTL
              HYHSHRPALGPIVVGVLKGLAEQYWGLGGEQLQVELLRGRDDGSDDHEVFRVSYPHQAALQHCGP
              PAASVDGNETAGSAASTASPAAAAARSFTLNPNDFYQLFPFHLLLDRSCRVVQAGAMLERLFPEL
              RGRSGVPLGEVFQLKHPHGPLDFDHLVSDIDNAFLLKARASGLELKGQMVAVPLPLHPGAGCPAT
              QEGLLFMGTARLAGLDDMRHHGIFISDIPHHDINRDYVLLAEQRQAEAQLQERLESLTRELKDAN
              SRLGEMAGWLDAERRRSDGLLYQMLPPEVASCLKNEERAPAQEHPEVTILFSDIVGFTEIASRSS
              PLEVCSLLDELYQRFDAAIEEYPQLYKVETIGDAYMVVCNVTVPCDDHADVLLEFALRMHEEASR
              VASSLGEPVRIRVGMHSGPVVAGVVGRKMPRFCLFGDTVNTASRMESHGEAGQIHISEACYSCLR
              SKERFEIRERGNITVKGKGTMRTYLLAPAPTSLEVLFQGPLEHHHHHHH*
```

S1B

```
Rn_alpha      TLLKQSSHCQEAERRGRLEDASILCLDKD-QDFLNVYYFFKRTTALLLFGIIKAAARILY 228
Rn_beta       ALHDHLATIY-----PGMRAPSFRCTDAEKGKGLILHYYSEREGLQDIVIGIIKTVAQQIH 160
Cr_Cyg11      DLHLHLSMSF-----PSMAAPAFKCTDVG-PTCLTLHYHSHRPALGPIVVGVLKGLAEQYW 158
Cr_Cyg12      DVHLHLGLMF-----PAMAVPAFECTDVG-PTCLKLHYHSHRPALGPIVVGVLKGLAEQYW 162
Cr_Cyg15      DLHLHLGMSF-----PAMAAPAFKCTDVG-PTCLTLHYHSHRPALGPIVVGVLKGLAEQYW 158
Cr_Cyg38      NLHLHLTMGF-----PSMVAPAFRVEEVT-GTSLLLHYHSVRPALWPIVVGVLKGASQKLF 157
Cr_Cyg56      VYHLHLSVTF-----KEMNFPAFNVSEVT-PQSLIFHYASQRPGLTRFAMGLIRGAARTLF 158
Cr_Cyg57      VYHLHLSMTF-----KELQFPAFNVSEVT-PDSLVFRYASTRPGLTRFAVGLLRGAARTLY 158
Consensus       H:H           :   P:   :      * . *YxSxR:     : G:::  :.
```

```
              *                                        ▲  *
Rn_alpha      DIVGFTAICSQC-SPLQVITMLNALYTRFDQQCG-ELDVYKVETIGDAYCVAGGLH-RES  541
Rn_beta       GIVGFNAFCSKH.GAMKIVNLLNDLYTRFDTLTD.NPFVYKVETVGDKYMTVSGLP-EPC  489
Cyg11         DIVGFTEIASRS-SPLEVCSLLDELYQRFDAAIEEYPQLYKVETIGDAYMVVCNVT-VPC  500
Cyg12         DIVGFTEIASRS-SPLEVCSLLDELYQRFDAAIEEYPQLYKVETIGDAYMVVCNVT-VPC  539
Cyg15         DIVGFTEIASRS-SPLEVCSLLDELYQRFDGAIEEYPQLYKVETIGDAYMVVCNVT-VPC  493
Cyg38         DMVGFTNTSAKS-TPEEVYVMLNEMYDKFDDVVDDFPDIYKLETIGDAYFLVCNLT.LPC  489
Cyg56         DIVGFTTISAGS-TAMEVCNMLDELYNEFDALLE.YKALYKVETIGDAYMLVANVT-EPC  712
Cyg57         DVVGFTTIAAGS-TAMEVCRMLDALYIEFDSLLE.YQSLYKVETIGDAYMLVANVT-KPC  671
Consensus     D:***.  .:      ::  :*: :* .**       :**:E*:*D *  . .:    .
```

```
                                        #                ▲▲     Δ
Rn_alpha      DTHAVQIALMALKMMELSNEVMSPHGEPIKMRIGLHSGSVFAGVVGVKMPRYCLFGNNVT  601
Rn_beta       IHHARSICHLALDMMEIAGQVQ-VDGESVQITIGIHTGEVVTGVIGQRMPRYCLFGNTVN  548
Cyg11         DDHADVLLEFALRMHEEASRVASSLGEPVRIRVGMHSGPVVAGVVGRKMPRFCLFGDTVN  560
Cyg12         DDHADVLLEFALRMHEEASRVASSLGEPVRIRVGMHSGPVVAGVVGRKMPRFCLFGDTVN  599
Cyg15         DDHADVLLEFALRMHEEASRVASSLGEPVRIRVGMHSGPVVAGVVGRKMPRFCLFGDTVN  553
Cyg38         DKHVDVVIDFAIRMQDAARQVKNAQGEPVQIRIGIHTGPVVGGVLGAKTPRFSIYGDTVN  549
Cyg56         VEHVDVMCDFALDMARVCSRVKTNRGQPLEIRVGIHTGSVVGGIVGRRMPRFHLFGDTVN  772
Cyg57         HDHVDVMLDFALDMHRVCSRVRTNRGQPIEIRIGVHSGSVVGGVVGRRMPRFTLFGDTVN  731
Consensus       *.  :  :*: *   . .*    *: :.:R:*:*:* *. *::* : *R:C::*:.*M
```

```
                 #                                      #
Rn_alpha      LANKFESCSVPRKINVSPTTYRLLKDCP----GFVFTPRSREELPPNFPSDIPGICHFLD  657
Rn_beta       LTSRTETTGEKGKINVSEYTYRCLMSPENSDPQFHLEHRG--PV------SMKGKKEPMQ  600
Cyg11         TASRMESHGEAGQIHISEACYSCLRSKE----RFEIRERG--NI------TVKGKGT-MR  607
Cyg12         TASRMESHGEAGQIHISEACYCCLRSKE----RFEIRERG--NI------TVKGKGT-MR  646
Cyg15         TASRMESHGEAGQIHISEACYSCLRSKE----RFEIRERG--NI------TVKGKGT-MR  600
Cyg38         VASRMESHGLPGKIHISGASYARIVDKH----KYRVRERG--NI------AVKGRGN-MT  596
Cyg56         TASRMESHGVPGSVHISGASRRFIRHPQ----RYLITERG--EI------SVKGKGL-ME  819
Cyg57         TASRMESHGVPGSVHMSGASRALLRDPS----QYLLTERG--EI------AVKGKGL-ME  778
Consensus       :.R *: .   .:::*      :         : .  *.   :       :K*    :
```

Figure B.1 – *Cr* sGC sequences and sequence alignments. A) The complete protein sequence of Cyg11615_His6 used in this study. B) Alignments of the Cr sGC H-NOX domains (top) and CAT domains (bottom) with Rn sGC α and β subunits. The critical residues are highlighted in green. Key residues: * – metal binding, Δ – guanine binding, ▲ – triphosphate binding, # – ribose binding.

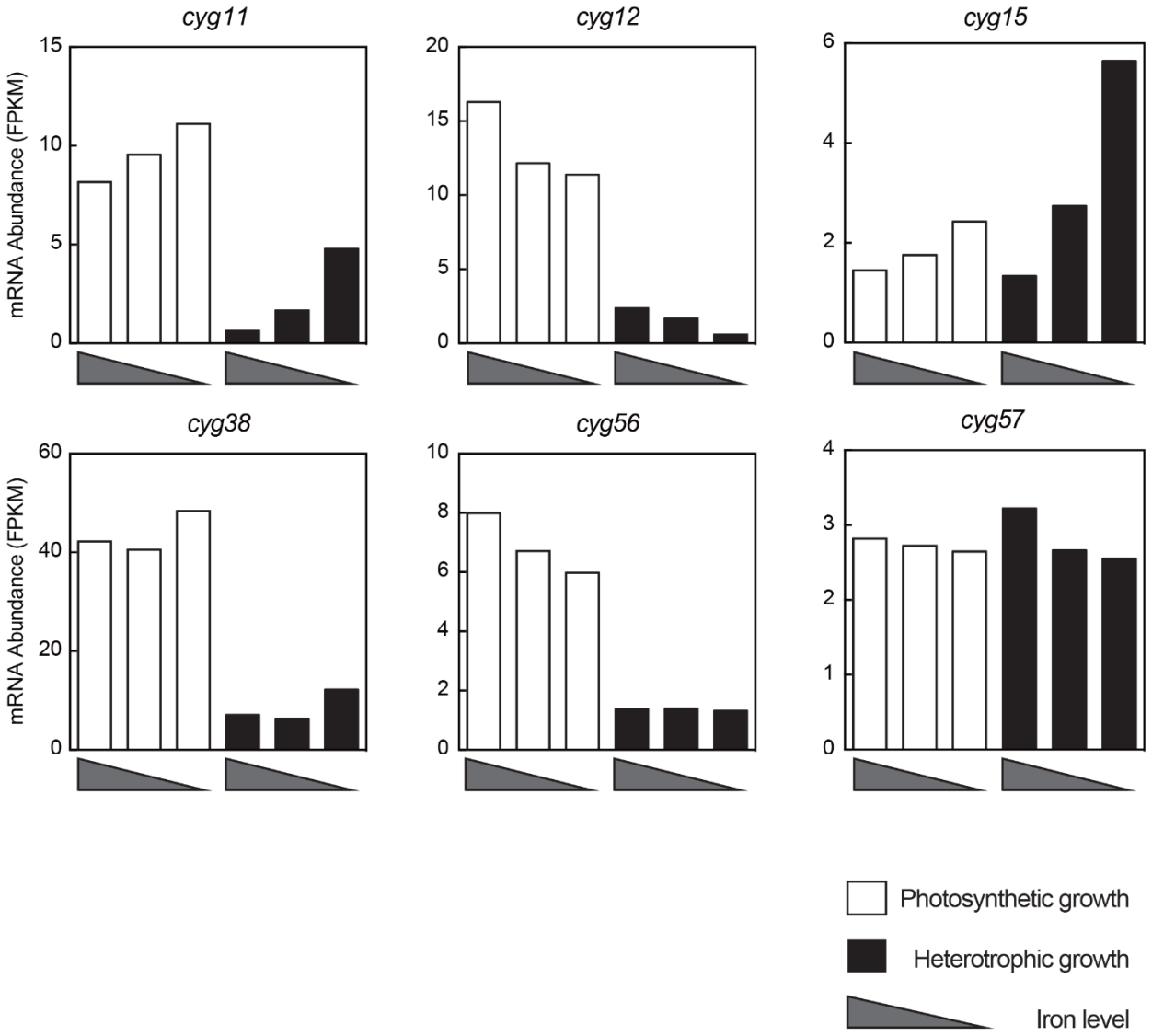


Figure B.2 – Transcriptomics of *Cr* sGCs in variable iron concentrations. The mRNA abundance for each *Cr* sGC is shown for three growth conditions with decreasing amounts of iron: replete, deficient, or limited. Heterotrophic growth signifies that *C. reinhardtii* was grown both in light and with acetate in the growth media as a carbon source, whereas photosynthetic growth indicates that C. reinhardtii was growth in light without acetate. Data collected from Phytozome (JGI) and the original data was published by Urzica et al.[1]

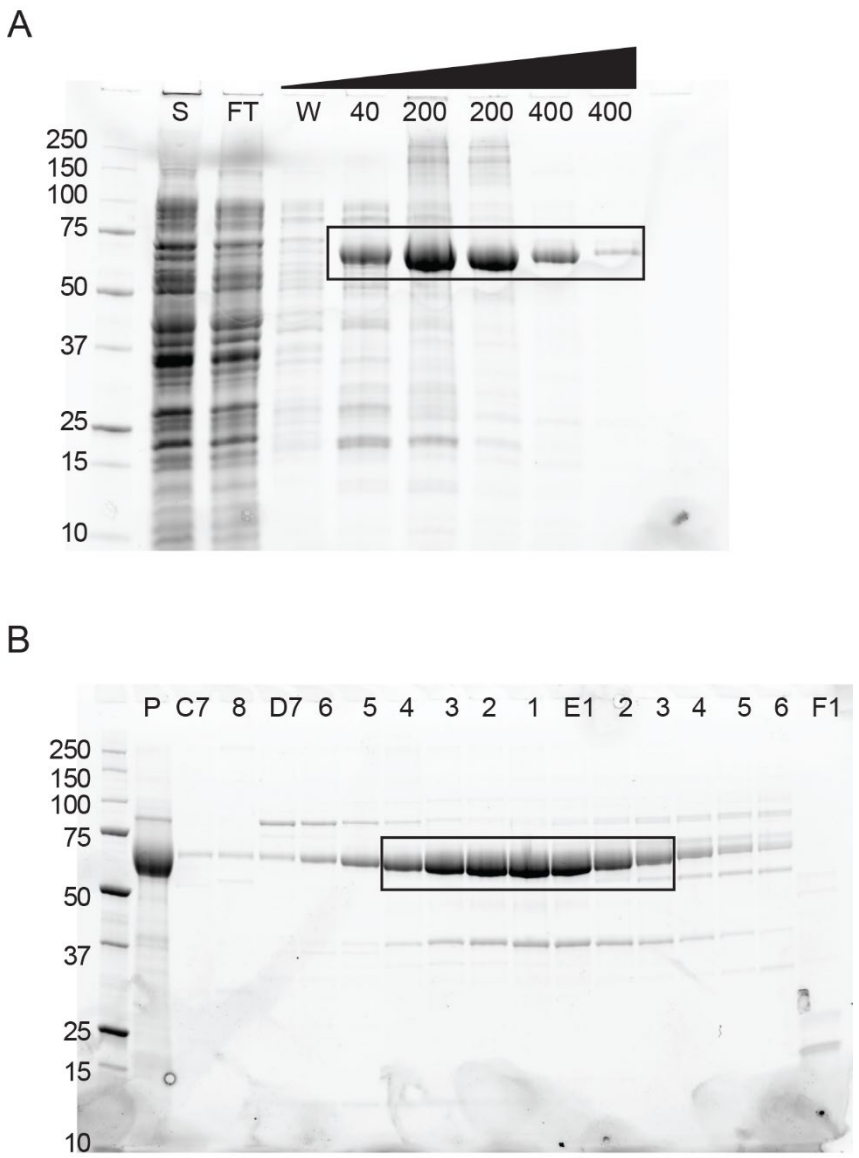


Figure B.3 – SDS-PAGE analysis of Cyg11 purification. A) SDS-PAGE gel of a typical immobilized metal affinity column purification of Cyg11 using His60 beads. S – soluble fraction, FT – flow through, W – wash; numbers indicate increasing concentrations of imidazole (in mM). B) SDS-PAGE gel of a representative size-exclusion chromatography purification of Cyg11. P – Pre; numbers indicate fractions collected from each of the three peaks present in the chromatogram. Boxes indicate fractions that were pooled and carried forward.

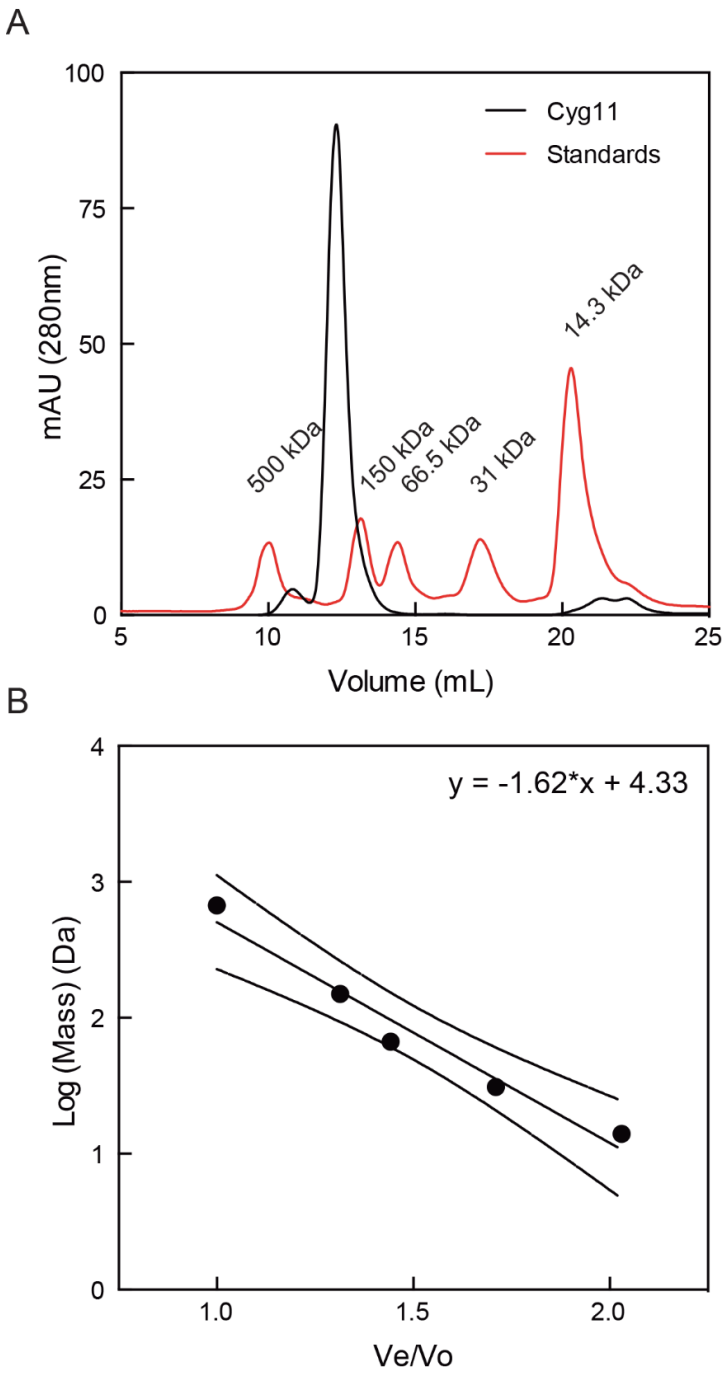


Figure B.4: – Cyg11 oligomeric state by analytical SEC A) The oligomeric state of Cyg11 was determined by analytical SEC, using five standards of known molecular weight: lysozyme (14.3 kDa), deoxyribonuclease I (31 kDa), bovine serum albumin (66.5 kDa), alcohol dehydrogenase (150 kDa), and thyroglobulin (500 kDa). The calculated molecular weight of Cyg11 using this method was approximately 200 kDa. B) Standard curve generated from the standards shown in panel A. 95% confidence curves are shown.

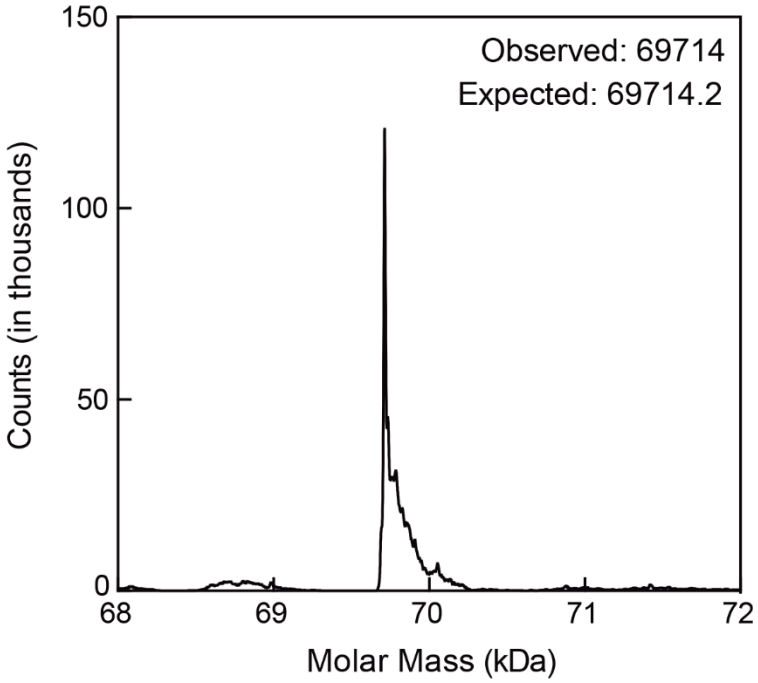


Figure B.5 – Intact protein mass spectrum of Cyg11. Purified Cyg11 was analyzed by intact protein mass spectrometry to validate the expression construct and support the results form native mass spectrometry analysis.

Ligation State	Protein	Soret Band	*β*	*α*
Unliganded				
	Rn $\alpha_1\beta_1$	431	555	
	So H-NOX	427	568	
	Cyg11	431	570	
NO				
	Rn $\alpha_1\beta_1$	398	537	572
	So H-NOX	398	544	572
	Cyg11	399	542	575
CO				
	Rn $\alpha_1\beta_1$	423	541	567
	So H-NOX	424	541	566
	Cyg11	424	537	569

Table B.1 – UV-visible Spectroscopy Peak Positions of sGCs (in nm). *Rn* sGC and *So* H-NOX peak positions obtained from the following references: [2,3]

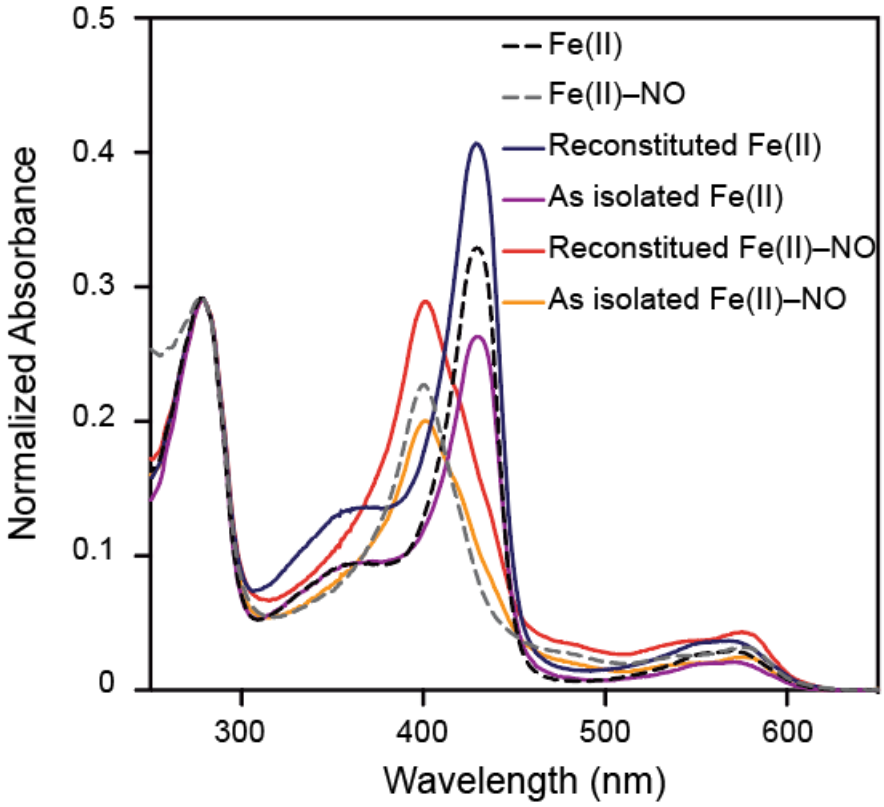


Figure B.6 – UV-visable spectra of reconstituted Cyg11. The typical ratio of 431 nm/280 nm ratio for Cyg11 is ~1.2. For this experiment, Cyg11 with a ratio of 0.9 was used. After reconstitution, the ratio of 431 nm/280 nm was 1.4. The incorporated heme is able to form a stable 5-coordinate ferrous nitrosyl complex, indicating that the heme is very likely present in the H-NOX domain.

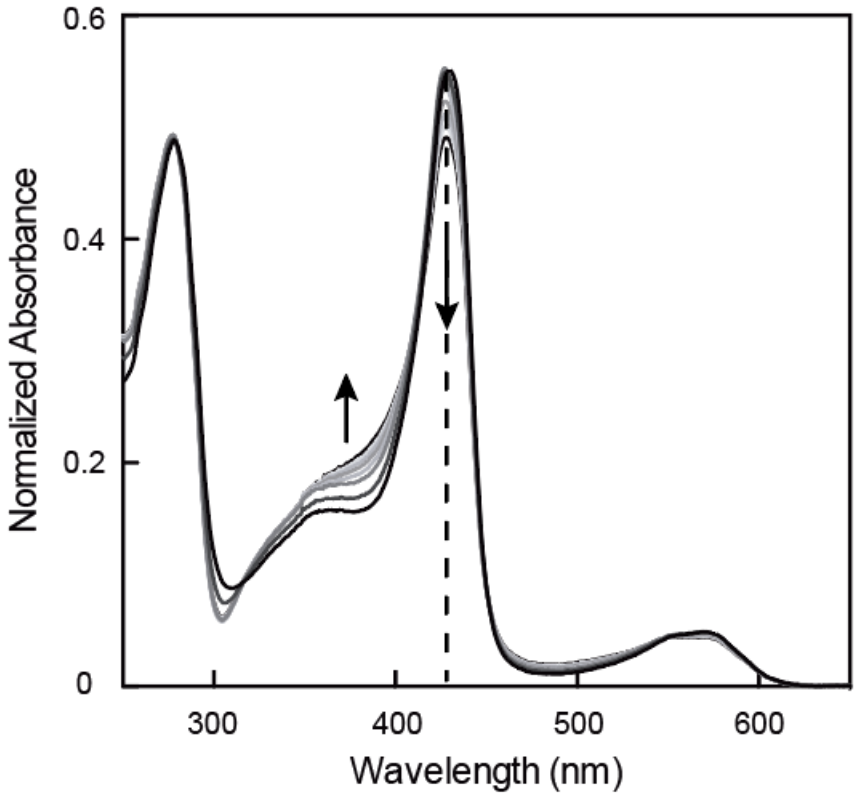


Figure B.7 – UV-visible spectra of Cyg11 Unliganded exposed to oxygen. Anaerobic, reduced Cyg11 unliganded was opened to ambient oxygen concentrations, and the UV/Vis absorbance spectra was recorded every hour for eight hours. The Soret band is centered at 431 nm (indicated by the dotted line) decreases in absorbance, most likeley due to oxidation and heme loss. However, the Soret does not shift to 416 m, as would be expected for a ferrous-oxy complex.[4]

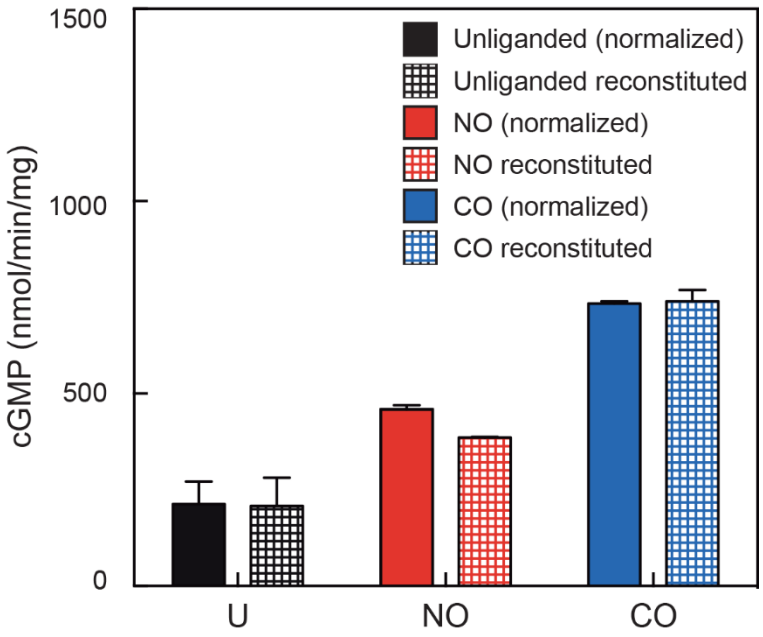


Figure B.8 – Effect of reconstitution on Cyg11 activity. Activity assays conducted with Cyg11 (431 nm/280 nm ratio = 0.9) were normalized to heme incorporation levels typically observed (431 nm/280 nm ratio = 1.2). This activity is comparable to Cyg11 that was reconstituted with hemin (431 nm/280 nm ratio = 1.4). Error bars represent standard deviation (n = 2).

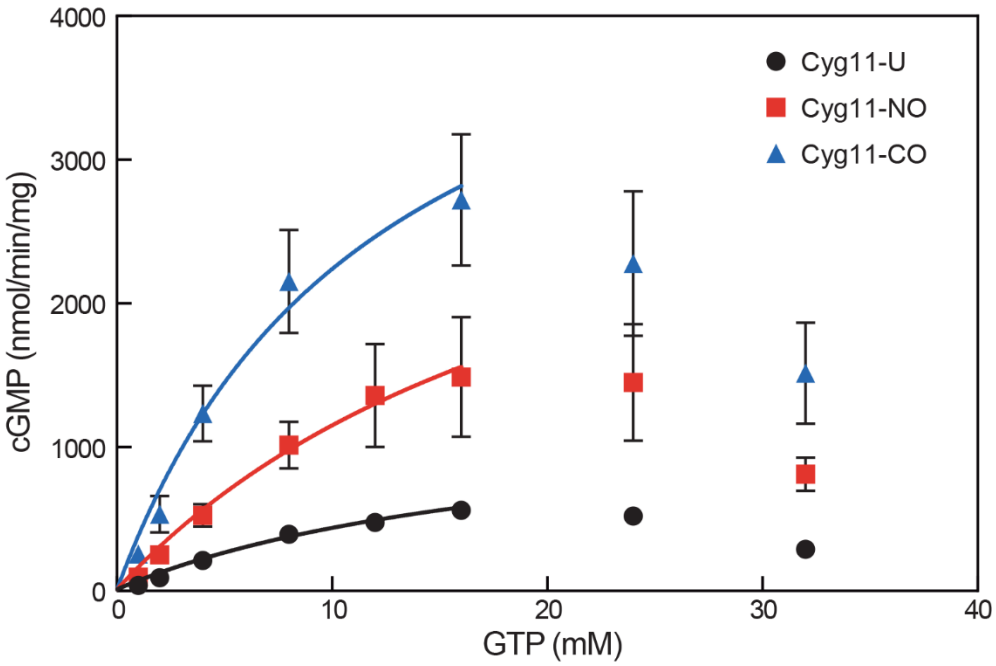


Figure B.9 – Michaelis–Menten kinetics of Cyg11 in different ligation states. Enzyme assays of Cyg11 with different ligation states was performed at 25 °C and pH 7.5 with varying GTP concentrations. Initial rates were calculated from each time course for 5–10% of substrate turnover using GraphPad Prism's nonlinear regression tools. Error bars represent standard deviation (n=3).

Appendix C

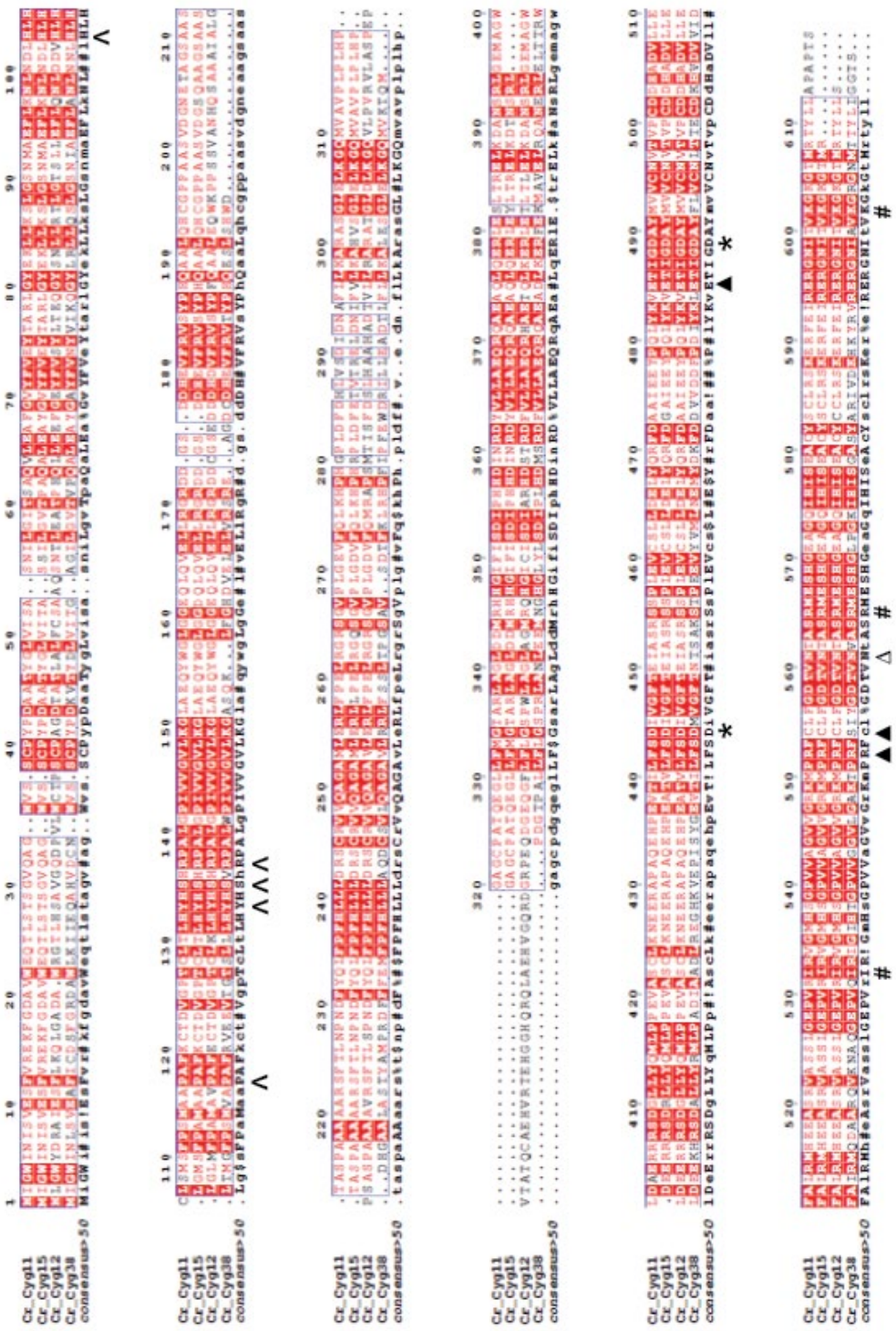


Figure C.1 – Sequence alignment of four *Cr* sGCs . Key residues: ^ – heme binding, * – metal binding, Δ – guanine binding, ▲ – triphosphate binding, # – ribose binding.

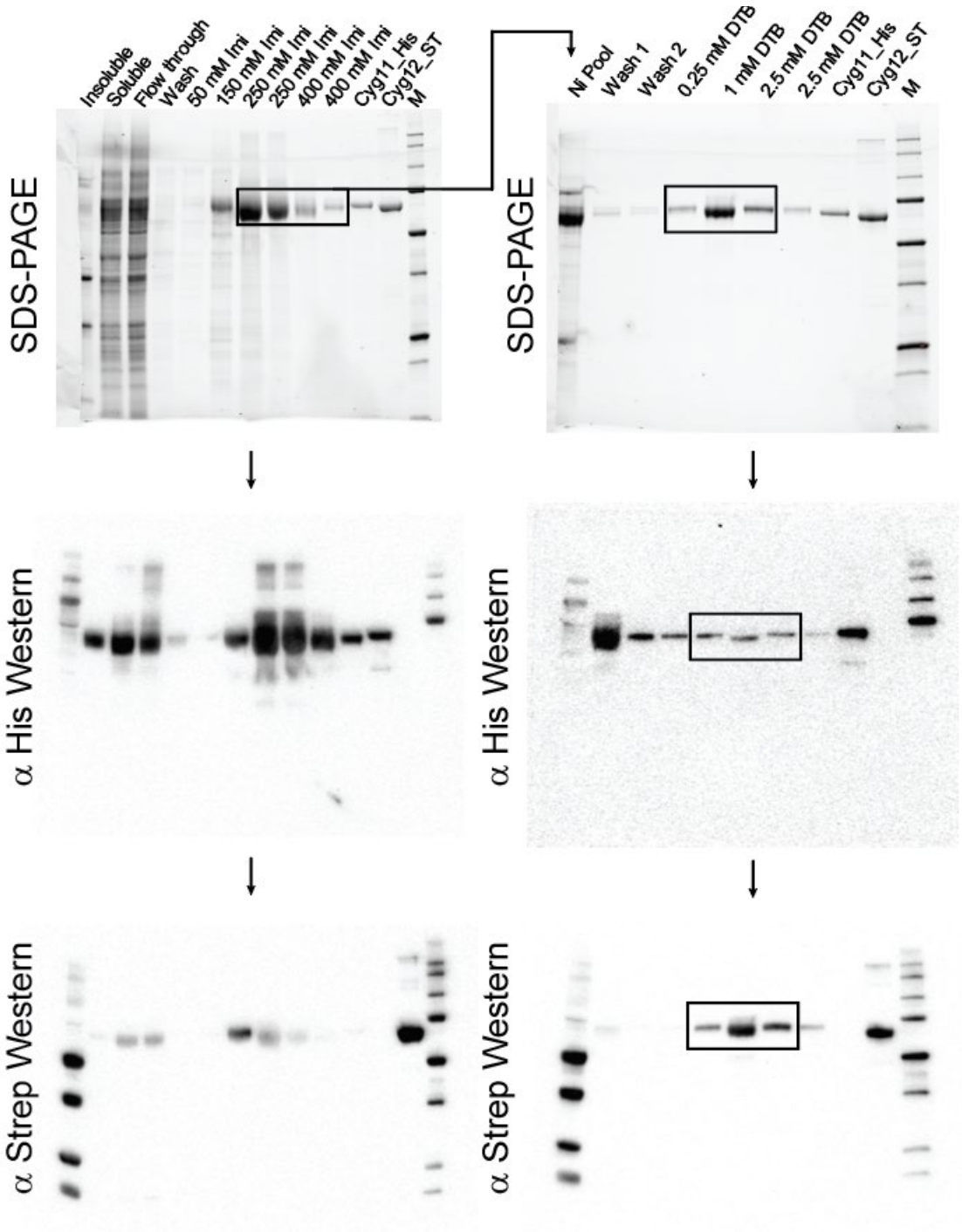


Figure C.2 – Purification gels and western blots of Cyg11-Cyg12 heterodimer. (Left top) SDS-PAGE of different fractions post Ni^{2+} IMAC. An α His western blot (middle) and a α Strep western are shown below. Boxed fractions were combined and assayed for activity and carried forward into the next column. (Right top) SDS-PAGE of different fractions post streptactin column. An α His western blot (middle) and a α Strep western are shown below. Boxed fractions were combined and frozen as the heterodimeric Cyg11-Cyg12.

Appendices References

(1) Urzica, E. I.; Casero, D.; Yamasaki, H.; Hsieh, S. I.; Adler, L. N.; Karpowicz, S. J.; Blaby-Haas, C. E.; Clarke, S. G.; Loo, J. A.; Pellegrini, M.; et al. Systems and *Trans*-System Level Analysis Identifies Conserved Iron Deficiency Responses in the Plant Lineage. *Plant Cell* **2012**, *24* (10), 3921–3948.

(2) Stone, J. R.; Marletta, M. A. Soluble Guanylate Cyclase from Bovine Lung: Activation with Nitric Oxide and Carbon Monoxide and Spectral Characterization of the Ferrous and Ferric States. *Biochemistry* **1994**, *33* (18), 5636–5640.

(3) Price, M. S.; Chao, L. Y.; Marletta, M. A. Shewanella Oneidensis MR-1 H-NOX Regulation of a Histidine Kinase By. *Biochemistry* **2007**, *46* (48), 13677–13683.

(4) Hespen, C. W.; Bruegger, J. J.; Phillips-Piro, C. M.; Marletta, M. A. Structural and Functional Evidence Indicates Selective Oxygen Signaling in Caldanaerobacter Subterraneus H-NOX. *ACS Chem. Biol.* **2016**, *11* (8), 2337–2346.